

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 17:49:28 ; Search time 70 Seconds
(without alignments)
1412.616 Million cell updates/sec

Title: US-10-029-347-2
Perfect score: 3218
Sequence: 1 MLAQPORLLFDGADLP...1THPDLGHQP?PKELISTP 625

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3098	96.3	947	11	US-09-965-621-18
2	1247	38.6	317	11	US-09-965-621-60
3	1084	33.7	290	11	US-09-965-621-61
4	825.5	25.7	1061	11	US-10-066-521-18
5	825	25.6	1035	11	US-09-965-621-24
6	800.5	24.9	1034	12	US-10-028-374-18
7	800.5	24.9	1034	12	US-10-132-967-5
8	800.5	24.9	1034	12	US-10-183-770-18
9	800.5	24.9	1034	14	US-10-127-516-5
10	800.5	24.9	1034	14	US-10-027-629-5
11	762	23.7	1429	10	US-09-986-617-2
12	762	23.7	1429	10	US-09-931-071-2
13	762	23.7	1429	12	US-10-028-374-15
14	762	23.7	1429	12	US-10-183-770-15
15	762	23.7	1429	15	US-10-028-392-11

16	762	23.7	1454	10	US-09-388-221-10
17	762	23.7	1473	10	US-09-388-221-2
18	760	23.6	1429	12	US-10-028-374-3
19	760	23.6	1429	12	US-10-183-770-3
20	716	22.2	1429	10	US-09-388-221-4
21	716	22.2	1424	10	US-09-388-221-12
22	716	22.2	1443	10	US-09-388-221-6
23	685	21.3	719	11	US-09-965-621-59
24	650	20.2	980	9	US-09-848-035-13
25	650	20.2	980	10	US-09-986-224-13
26	641	19.9	449	12	US-10-028-374-2
27	641	19.9	449	12	US-10-183-770-2
28	634	19.7	1049	12	US-10-233-663-42
29	619.5	19.3	1033	12	US-10-132-967-2
30	619.5	19.3	1033	14	US-10-127-516-2
31	619.5	19.3	1033	14	US-10-027-629-2
32	616.5	19.2	1062	12	US-10-239-663-43
33	598.5	18.6	858	9	US-09-848-035-8
34	598.5	18.6	858	10	US-09-986-224-8
35	598.5	18.6	994	11	US-09-965-621-16
36	598.5	18.6	994	15	US-10-066-521-24
37	597.5	18.6	1016	10	US-09-986-224-19
38	586.5	18.2	896	15	US-10-066-521-22
39	567.5	17.6	514	12	US-10-028-374-14
40	567.5	17.6	514	12	US-10-183-770-14
41	567.5	17.6	514	15	US-10-028-392-10
42	511.5	15.9	1143	15	US-10-216-645-4
43	511.5	15.9	1162	15	US-10-216-645-2
44	511.5	15.9	1344	15	US-10-066-521-6
45	501.5	15.6	821	9	US-09-848-035-2

ALIGNMENTS

RESULT 1
US-09-965-621-18
; Sequence 18, Application US/09965621
; Publication No. US20030077699A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stenlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/09/965,621
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.3
; SEQ ID NO 18
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-965-621-18

Query Match 96.3%; Score 3098; DB 11; Length 947;
Best Local Similarity 90.0%; Pred. No. 1e-286;
Matches 613; Conservative 1; Mismatches 11; Indels 56; Gaps 2;
QY 1 MLAQPORLLFDGADLPALGGPEAAPCTDPFEAASGARVLGGLSKALLPTALLVTT 60
Db 267 MLAQPORLLFDGADLPALGGPEAAPCTDPFEAASGARVLGGLSKALLPTALLVTT 326
QY 61 RAAAPCRLOGRLCSQCAEVRGFSKDKKKYFYKFFRDERAERAYRVKENETLFCALCF 120
Db 327 RAAAPGRLOGRLCSQCAEVRGFSKDKKKYFYKFFRDERAERAYRVKENETLFCALCF 386

60/367,367
Converted to

```
QY 121 VPFVCMIVCTVLRQOELGRDLSRTSKTTTSVYLFITSVLSAPVADGPRQGLRNLC 180
DB 397 VPFVCMIVCTVLRQOELGRDLSRTSKTTTSVYLFITSVLSAPVADGPRQGLRNLC 446
QY 181 RLAREGVLGRRQAQFAKELEOLELRGSKVQTLFLSKKELPGVLETEVTYQFIDQSFOEFL 240
DB 447 RLAREGVLGRRQAQFAKELEOLELRGSKVQTLFLSKKELPGVLETEVTYQFIDQSFOEFL 506
QY 241 AALSYLEDDGVPRTAAGGVTLLRQDAQPHSHLVLTTRFLGLLSAERNRDIERHFGCM 300
DB 507 AALSYLEDDGVPRTAAGGVTLLRQDAQPHSHLVLTTRFLGLLSAERNRDIERHFGCM 566
QY 301 VSERVVKQEAALRWYQGGQCGFCGVAPEVTGAKGLEDTESPEEEEGEENYPLELLYCLY 360
DB 567 VSERVVKQEAALRWYQGGQCGFCGVAPEVTGAKGLEDTESPEEEEGEENYPLELLYCLY 626
QY 361 ETOEDAFVQOALCRFPPELALQVRFCRMDVAVLSYCVRCPCPAGQALRLISCLVAQAQKK 420
DB 627 ETOEDAFVQOALCRFPPELALQVRFCRMDVAVLSYCVRCPCPAGQALRLISCLVAQAQKK 686
QY 421 KXSLGKRLQASLGGSSQGT----- 443
DB 687 KXSLGKRLQASLGGSSQGT----- 746
QY 441 TKQLPASLL-----HPLFCAMTDPLCHLSSJTLGSHCKLPDCAVCRDLSEA 464
DB 747 IOPLFRLGLASVERMNVTLGAGCGDPKPTKNTDPLCHLSSJTLGSHCKLPDCAVCRDLSEA 806
QY 485 LRAAPALTEGLLHNRLSEAGLRMSGLAWPQCRVQTVRVVLPDQORGLQVVGMLRQS 544
DB 807 LRAAPALTEGLLHNRLSEAGLRMSGLAWPQCRVQTVRVVLPDQORGLQVVGMLRQS 866
QY 545 PALTLDLGGCOLPAPMWYTLCAVLQHQCGGLQTLASVELSEQLQLOAVKAKPDL 604
DB 867 PALTLDLGGCOLPAPMWYTLCAVLQHQCGGLQTLASVELSEQLQLOAVKAKPDL 926
QY 605 VITHPALDGHPPPKELISTF 625
DB 927 VITHPALDGHPPPKELISTF 947

RESULT 2
US-09-965-621-60
; Sequence 60, Application US/09965621
; Publication No. US2003007699A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-621-60

Query Match 38.8%; Score 1247; DB 11; Length 317;
Best Local Similarity 100.0%; Pred. No. 1.7e-110;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MLAPQORLLFLIDGADLPPALGGPEAAPCTDPPRAASGARVILGGLSKALLPTALLVTT 60
DB 72 MLAPQORLLFLIDGADLPPALGGPEAAPCTDPPRAASGARVILGGLSKALLPTALLVTT 131
QY 61 RAAAPGRLOGRLCSPQCAEVGRFSDKDKKYFYKFFDERRAERAYRFVKENETLFCALCF 120
DB 132 RAAAPGRLOGRLCSPQCAEVGRFSDKDKKYFYKFFDERRAERAYRFVKENETLFCALCF 191
QY 121 VPFVCMIVCTVLRQOELGRDLSRTSKTTTSVYLFITSVLSAPVADGPRQGLRNLC 180
DB 192 VPFVCMIVCTVLRQOELGRDLSRTSKTTTSVYLFITSVLSAPVADGPRQGLRNLC 251
QY 181 RLAREGVLGRRQAQFAKELEOLELRGSKVQTLFLSKKELPGVLETEVTYQFIDQSFOEFL 240
DB 252 RLAREGVLGRRQAQFAKELEOLELRGSKVQTLFLSKKELPGVLETEVTYQFIDQSFOEFL 311
QY 241 AALSYL 246
DB 312 AALSYL 317

RESULT 3
US-09-965-621-61
; Sequence 61, Application US/09965621
; Publication No. US2003007699A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-621-61

Query Match 33.7%; Score 1084; DB 11; Length 290;
Best Local Similarity 76.6%; Pred. No. 5.6e-95;
Matches 222; Conservative 1; Mismatches 11; Indels 56; Gaps 2;
```

```
QY 392 VLSYCVRCPCPAGQALRLISCLRLVAQAQKKKSLGKRLQASLGGSSQGT----- 440
DB 1 VLSYCVRCPCPAGQALRLISCLRLVAQAQKKKSLGKRLQASLGGSSQGT----- 60
QY 441 -----TKQLPASLL-----HPLFQA 455
DB 61 CCDICPTPPDPRLLQKAFARVPLNTAPIQPLRGLASVERMNVTLGAGCGDPKTHA 120
QY 456 MTDPLCHLSSJTLGSHCKLPDAVCRDLSEALRAAPALTEGLLHNRLSEAGLRMSGLAW 515
DB 121 MTDPLCHLSSJTLGSHCKLPDAVCRDLSEALRAAPALTEGLLHNRLSEAGLRMSGLAW 180
QY 516 PQCRVQTVRVQLPDQORGLQVVGMLRQSPALTTLDLGGCOLPAPMWYTLCAVLQHQCGG 575
DB 181 PQCRVQTVRVQLPDQORGLQVVGMLRQSPALTTLDLGGCOLPAPMWYTLCAVLQHQCGG 240
QY 576 LOTLSLASVELSEQLQLOAVKAKPDLVITHPALDGHPPPKELISTF 625
DB 241 LOTLSLASVELSEQLQLOAVKAKPDLVITHPALDGHPPPKELISTF 290
```



```

RESULT 4
US-10-066-521-i8
; Sequence 18, Application US/10066521
; Publication No. US20030027757A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/JRR
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-334001
; CURRENT APPLICATION NUMBER: US/10/066,521
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/318,645
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/265,231
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1061
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-066-521-18

Query Match      25.7%; Score 825.5; DB 15; Length 1061;
Best Local Similarity 32.6%; Pred. No. 2.1e-69;
Matches 219; Conservative 104; Mismatches 245; Indels 103; Gaps 15;

QY      1  MLAGPQRLLFLDCADEL-PALGQPEAACTDPPEASGARVLGSLKALLPTALLLVT 59
DB      282 LIRVPERLLFIIDGFDLKHSPHDQGPWCLEWKEKRTPELLNSLIRKKLPELSLIT 341

QY      60  TRAAAPGRLOGRLCSPQCAEYRGFSDDKKKYFYKFFRDERRAERAYRFVKENETLFC 119
DB      342 TRPTALEKLHRLLEHPRHVEILGFSAEERKEYFYKYFHNAEQAGQVFNVRDNEPLFTMC 401

QY      120 FVPEVCHTVCTVLPQOELGRDLRSKTTTSVLLFTITSVLSSAPVADGRLO--GDLR 177
DB      402 FVPLVCVVVCTCLQQQEGGGGLRQTSRTTTAVYMLYLLSLMQPKPGA--PRLOPPNQR 459

QY      178 NLCLRAEGVLGRRAQFAEKELEOLELRGSKVQTLFLSKKELPGVLETVYQYFIDQSFO 237
DB      460 GLCSLAADGLWNQKLFEEQDLRKHGLDGEDV-SAFLNWNTFKDINCERYYSFIHLSFO 518

QY      238 EFLAALSYLLEDGGVPRTPAAGVG-----TLRGDAQPHSHLVLTTRFIFGLLSAERM 290
DB      519 EFFAAMYILDEG-----EGCAGPDQDVTRLLTEYAFSEFSFALTSRFLFGLNSETR 572

QY      291 RDIERHFGCMYSERYKQEAALWQOGQCGQVAPETVEGAKGLECTEPEEEEEGEEN 350
DB      573 SHLEKSLCWKYSPIHKMOLLQWISKAO-----SDGSLTQGS--- 610

QY      351 YPLELLYCLYETQBDADFVRQALCRFPPELALQVRFCRMDVAVLGYCYVRCPCPAGALRLIS 410
DB      611 --LEFFSCLYEIEEFTQQAALSHFQVIVWSNIA-SKMEHWVSFCLKRCRKSQVLYHLG 667

QY      411 CRLVAAEKKKSKGLKRLQASLGGGS-----SQGTTKQIPASL----- 448
DB      668 ATYGADGE-----DRARCSAGAHRTLLVQLPFTVLLDDAYSEHLAAALCTNPNIELS 720

QY      449 -----LHPLFQAMTDPLCHLSLSTLSHCKLPDAVCRDLSEALRAAPALTEGLCHN 499
DB      721 YRNALSGRVKLLCOGLRHPNCKQLNLRKRCRISSACEDLSAALIANKNLTRMDLSGN 780

QY      500 RLSBAGLRMLSEGLAWPOCRVQTVRVVQYDPQRG--LQVLQMLRQSPALTTTLDLSGGQCP 558
DB      781 GVGFPGWMLLCEGLRHHPQCRLCMTQLRCKQLESQACQEMASVLTGNPHLVELD.TGNALE 840

QY      559 APMTYLYCAVLQHOCCGGLQTL-----SLASVELSEQSLOELQAVKRAKPDIV 605
DB      841 DLGRLRLCQGRHPVCRVLRITLWKICRLTAACADELASTLSVNSQSLRELDLSLNLGDLG 900

```

```

Cy 606 I-----THP 609
:
Db 901 VLLICEGLRH? 911

RESULT 5
US-09-965-621-24
: Sequence 24, Application US/09965621
: Publication No. US20030077699A1
: GENERAL INFORMATION:
: APPLICANT: Reed, John C.
: APPLICANT: Godzik, Adam
: APPLICANT: Chu, Zhi-Liang
: APPLICANT: Pawlowski, Krzysztof
: APPLICANT: Fiorentino, Loredana
: APPLICANT: Ariza, Maria Eugenia
: APPLICANT: Stehlik, Christian
: TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
: TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
: FILE REFERENCE: P-LJ 4816
: CURRENT APPLICATION NUMBER: US/09/965.621
: CURRENT FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: US 09/671,760
: PRIOR FILING DATE: 2000-09-26
: NUMBER OF SEQ ID NOS: 64
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 24
: LENGTH: 1035
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-965-621-24

Query Match: 25.6%; Score 825; DB 11; Length 1035;
Best Local Similarity 32.6%; Pred. No. 2,3e-69;
Matches 219; Conservative 104; Mismatches 245; Indels 104; Gaps 15

Cy 1 MLAQPORLLFILDGADL-PALGGPEAAPCTDPEAASGARVLGGLLSKALLPTALLVT 59
Db 255 LIRVPERLLFIIDGDLKPSFHDQPGPCLWCBEKRPTELLNLSLRKLLPELSLIT 314
Cy 60 TRAAAPORLOGRLCSPOCAEVGFSDKDKKYYFKFRDRERAERAVFVKENTLFLALC 119
Db 315 TRPTALEKHLRLLEHPRHVEILGFSEAEKXEYFYKYFHNAEQAQGVNYVRDNEFLFTMC 374
Cy 120 FVPFVCWIVCTLRQQLELGRDLRSKTTTSVYLLFITSVLSSAPVADGRLQ--GDLR 177
Db 375 FVPLVCWVCTLCQQLEGGGLLRQTSRTTAVYMLYLLSLMQPKPGA--PRLOPPNQ 432
Cy 178 NLCLRLREGVLGRRAQFAEKLEQLGRLGSKVQTLFLSKKELPGVLETVYQFIDQSFQ 237
Db 433 GLCSLAADGLWNKILPEEQDLRKHLDGEDV-SAFLLMN:FQXDINCERYYSFIHLSFQ 491
Cy 238 SEUALSLYLLDGGVPTAAGVG-----TLIRGDAQPHSHLVLTRELFGLLSAERM 290
Db 492 EFTAAMYIILDEG-----EGGAGPODVTRLTLEYAFSERSFALTSRFLUGLLNEET 545
Cy 291 RD:ERHFGCMVSERVKOEALRWVGQGGCGPVAPEVTEGAKLEDTEEPPEEEGEBEPN 350
Db 546 SHLEKSLCWKVSPIKMDILQWIOSKAQ-----SDGSTLQGS----583
Cy 351 YPIELLYCLYETQEDAFVROALCRFPPELALQVRFCRMDVAVLSYCVRCPCPAGQALRLIS 410
Db 584 --LEFFSCLEYIOEEEFIQQALSHFQIVIVVSNIA-SKWEHVVSSFCLKRCRQAQVLYLYG 640
Cy 411 CRLVAQEKKKKSLGKRLQSLGGGS-----SOGTKQLPASL-----448
Db 641 ATYSADGE-----DRARCSAGHTLLVQLRPRTVLLDAYSEHLAAALCTNPNL:ELS 693
Cy 449 -----LHPLFOAMTDPCHLSLSLTSHCKLPDVCEDLSEALRAAPALTELGLLH 498
Db 694 LYRNALGSRGKULLCQGLRHPNCKLQNLRLKRCRISACEDUSAALTIANKNLTRMDLSG 753
Cy 499 NRLSEAGRLMSLSEGLAWPQCRVQTVRLQDPDPORG-LQYLVGLMRQSPALVTLLDSGCQL 557

```

```

Db      754 NGVGFPGWMLLCEGLRHPQCRLOMQLKCKQESACQEMASVLTGTHLDELDTGVAL 813
QY      558 PAPMTYTLCAVLQHQCGGLOTL-----SLASVELSEQLQELQAVKRAKPDIL 604
Db      814 EDLGLRLLCQGLRHPVCRURTILWLKICRUTAAACDELASTLSVNGSLRELDLSNELGDL 873
QY      605 V-----THP 609
Db      874 GVLLCEGLRHP 885

RESULT 6
US-10-028-374-18
; Sequence 18, Application US/10028374
; Publication No. US20030143706A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE REFERENCE: D0067NP
; CURRENT FILING DATE: 2001-12-20
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-374-18

Query Match      24.9%; Score 800.5; DB 12; Length 1034;
Best Local Similarity 28.4%; Pred. No. 5.1e-67;
Matches 219; Conservative 117; Mismatches 241; Indels 193; Gaps 17;

QY      1 MLAQQRLLFIIDGADLPALGGPEAAP-CTDPFEAASGARVVGGLLSKALLPTALLIVT 59
Db      288 IVRKPSRIILFMDGFDDELQAGDEHIGPLCTDQKAERGDIILSSLRKLLPEASLLIT 347
QY      60 TRAAAPGRLLQGRLLCSQCAEVRGFSKDKKKYFKFFRDEREAAERAYRVFVKNETLFAIC 119
Db      348 TREVALEKLOHLLDHPRHVEILGFSEAKREYFKYFSDQAQARAAPSLIQENEVLFTMC 407
QY      120 FVPFVCMIVCTVLRQOLEGRDLRSKTTTSVYLLFTSVLSAPVADGPRLOGLRLNL 179
Db      408 FIPLVCMIVCTGLKQOMESGKSLAQTSTKTITAVYVFFLSLLQPRGSGQEHGCAHLWGL 467
QY      180 CRLAREGLVGRRAQFAEKELEQLRSGSKVOTLFLSKKELPGVLETVTYQIDQSQFQEF 239
Db      468 CSLAAGDIWNQKILFESDLRNHGLQKADV-SAFLRMNLFQKEVDCFKFYSFIHMTQFQEF 526
QY      240 LAALSYLEDDGGVPRTAAGG-----VGTLLRGDAQ- PSHSLVLTTRFLFGLLSAER 289
Db      527 FAAMYLYLBEKEGRNVPGSRLLKPSRDVTVLLENYGFKEGYLIFVVRFLFGLVNGQER 586
QY      290 MRDIERHFCMYSERVKQALRWVQGGCGCPGVAPEVTEGAKGLEDETEPEEBEEGEEP 349
Db      587 TSYLEKKLSCKISQIIRLELLKMI-----EVKAKAKKLO--IQPSQ----- 625
QY      350 NYPLELLYCLYETQEDAFVQALCRPELALQVRFCRMDVAVLSVCVRCPPAGQALRLI 409
Db      626 ---LELFYCLYEQEEDFVQRANDYFPKTEIN--LSTRMDHVMVSSFCIENCHRVESLSLG 680
QY      410 SCLVAAQKKKKSLGKRLQ-----ASLGGSSQG-TTKQLPASLLHPLF-----QA 455
Db      681 FCHNMPKEEEEEKRHLDMYQCVLPSSSHAACSHGLVNSHLTSSFCRGLFSVLSTSGS 740
QY      456 MT-----DP----- 459
Db      741 LTJELDSLNSLDGPGMRVLCITLOHPGNCNIRRLWLGRCGLSGCECCFDISLVLSNQKQVE 800
```

```

QY      460 -----LCHLSSLTLSHCKLPDAVCRDLSEALRAAPALTELGL 496
Db      801 LDLSDNALGDFGIRLLCVGLKHLKMLWVSCCLTSACCDLASVLTSHSLTRYV 86C
QY      497 LHNRLSEAGRLMSEGLAWPQCRVQTVRV----- 525
Db      861 GENALGDSGVAIILCEKAKNPQCNLOKGLVNSGLTSCCSALSVLSTNQNLTHLYLRGN 920
QY      526 QLPDPQRGLQYLV-GMLRQSPALTTTLDLSCQCLPAPM----- 561
Db      921 TLGD--KGIKLLCEGLLHPDKLVLELDNCNLTSCHCCWDLSTLTSSQSLRKLSLGNND 978
QY      562 -----VTYLCVAVLQHQCGLOTLSLASVELSEQLQELQAVKRAKPDILVI 606
Db      979 LGLGVNMFCEVLKQOQSCLLQNLGLSEMYFYETKSALETLOEEKPELTV 1028

RESULT 7
US-10-132-967-5
; Sequence 5, Application US/10132967
; Publication No. US20030170841A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN PROTEIN FAMILY AND USES THE
; FILE REFERENCE: 07334-311001
; CURRENT FILING DATE: 2002-04-26
; PRIOR FILING DATE: 2000-09-01
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-967-5

Query Match      24.9%; Score 800.5; DB 12; Length 1034;
Best Local Similarity 28.4%; Pred. No. 5.1e-67;
Matches 219; Conservative 117; Mismatches 241; Indels 193; Gaps 17;

QY      1 MLAQQRLLFIIDGADLPALGGPEAAP-CTDPFEAASGARVVGGLLSKALLPTALLIVT 59
Db      288 IVRKPSRIILFMDGFDDELQAGDEHIGPLCTDQKAERGDIILSSLRKLLPEASLLIT 347
QY      60 TRAAAPGRLLQGRLLCSQCAEVRGFSKDKKKYFKFFRDEREAAERAYRVFVKNETLFAIC 119
Db      348 TRVALEKLOHLLDHPRHVEILGFSEAKREYFKYFSDQAQARAAPSLIQENEVLFTMC 407
QY      120 FVPFVCMIVCTVLRQOLEGRDLRSKTTTSVYLLFTSVLSAPVADGPRLOGLRLNL 179
Db      408 FIPLVCMIVCTGLKQOMESGKSLAQTSTKTITAVYVFFLSLLQPRGSGQEHGCAHLWGL 467
QY      180 CRLAREGLVGRRAQFAEKELEQLRSGSKVOTLFLSKKELPGVLETVTYQIDQSQFQEF 239
Db      468 CSLAAGDIWNQKILFESDLRNHGLQKADV-SAFLRMNLFQKEVDCFKFYSFIHMTQFQEF 526
QY      240 LAALSYLEDDGGVPRTAAGG-----VGTLLRGDAQ- PSHSLVLTTRFLFGLLSAER 289
Db      527 FAAMYLYLBEKEGRNVPGSRLLKPSRDVTVLLENYGFKEGYLIFVVRFLFGLVNGQER 586
QY      290 MRDIERHFCMYSERVKQALRWVQGGCGCPGVAPEVTEGAKGLEDETEPEEBEEGEEP 349
Db      587 TSYLEKKLSCKISQIIRLELLKMI-----EVKAKAKKLO--IQPSQ----- 625
QY      350 NYPLELLYCLYETQEDAFVQALCRPELALQVRFCRMDVAVLSVCVRCPPAGQALRLI 409
Db      626 ---LELFYCLYEQEEDFVQRANDYFPKTEIN--LSTRMDHVMVSSFCIENCHRVESLSLG 680
QY      410 SCLVAAQKKKKSLGKRLQ-----ASLGGSSQG-TTKQLPASLLHPLF-----QA 455
```

```
Db 681 FLHNPKEEBEKEGRHLDVQCVPVLPSSSHACSHGLVNSHLTSSFCRGLFSVLSTQS 740
Qy 456 MT-----DP-----
Db 741 LTELDSNLSLGDPGMRVLCETLQHPGNCIRRLWLGRGCLSHCECCFDISLVSSNQKIVE 800
Qy 460 -----LCHLSLTLSHCKLPDAVCRDLSEALRAAPALTEGL 496
Db 801 LLSNALGDFGIRLLCVGLKHLNCKLKLWLVSCLLTSACODLASVLSTSHSLRLYV 860
Qy 497 LHNRLSEAGRLMSEGLAMPQCRVQTVRV----- 525
Db 861 GENALGDSGVALICEKAKNPQCNLQKGLVNSGLTSVCCSALSSVLSTNQNLTHLYLGN 920
Qy 526 QLPDPORGLOYLIV-GMLROSPALTTLDLSGCCOLPAPM----- 561
Db 921 TLGD--KGIKLCEGILLHPDCKLQVLELDNCLNLTSHCCWDLSTLTSSQSLRKLSLGNND 978
Qy 562 -----VTYLCVAVLQHOGCGLOTLASVELSSEQLQAVKRAKPDVVI 606
Db 979 LGDLGWMFCEVLKQOSCLQNLGLSEMFTYETKSALETLOEKEPELTV 1028

RESULT 8
US-10-183-770-18
; Sequence 18, Application US/10183770
; Publication No. US20030180812A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLIRREB1
; FILE REFERENCE: D0067A CIP
; CURRENT APPLICATION NUMBER: US/10/183,770
; PRIOR FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/328,374
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-183-770-18

Query Match 24.9%; Score 800.5; DB 12; Length 1034;
Best Local Similarity 28.4%; Pred. No. 5.1e-67;
Matches 219; Conservative 117; Mismatches 241; Indels 193; Gaps 17;

Qy 1 MLAQFORLLFIIDGADLPALGPPEAAP-CTOPFEAASGARVLGGLLSKALPTALLVLT 59
Db 288 IVKPSRILFLMDGFDLOGAFDEHIGPLCTDQKAERGDILLSSLRKKLPEASLLIT 347
Qy 60 TRAAAPGRLOGRLCSQCAEVGFSKDKKKYFYKFFRDRERAERAYRFVKNETLFLALC 119
Db 348 TRVALEKQLHLDHPRHVEILGFSSEAKKEYFFKYFSDQAQAAAFSLIOENEVLFTMC 407
Qy 120 FVPFVCWIVCTVLRQOLEGRDLRSRTKTTTSVYLLFITSVLSSAPVADGPRLOGLDRLN 179
Db 408 FIPLVCMIVCTGLKQMESGKSLAQTSKTTITVYVFFLSLLQPRGGSGEHLCAHLWGL 467

Qy 180 CRLAREGVLRRAQFAKEKELEQLRGSKVQTLFASKELPGVLETEVTVYQIFDOSQEF 239
Db 468 CSLAADGIWNOKILFEESDLRNHGLQADY-SAFRLMNLFOKEVCEKFPYSFIHTFOEF 526
Qy 240 LAALSILLEDGGVPRTAAG-----VGTLLRGDAQ-PHSHLVLTTRFLFGLLSAER 289
Db 527 FAAMYVLLLEEKGRTNVPQSKLPSRDTVLLNRYGKFXEYGLFVVRFLFGLVNOER 586
Qy 290 MRDIERHFGCWSEKVKQALRWVQGGGCGCPVAPENVTEGAKGLEDTEPEPEEEEGEEP 349
Db 587 TSYZEKMLSKISQIRLELLRWI-----EVKAKKKLQ--IQPSQ----- 625
```

```
Qy 350 NYPLELLYCLYETQEDAFVVRQALCRFPPELALQVRFCMDVAVLSYCVRCPCPAGALRLI 409
Db 626 ---LELFYCYEMOEEDFVORAMDFPKIEIN--LSTRMDHMVSSFCIENCHRVSLSLG 680
Qy 410 SCRLVAQKKKKSLGKRLQ-----ASLGGSSQG-TTKOLPASLHLPLF-----QA 455
Db 681 FLHNPKEEBEKEGRHLDVQCVPVLPSSSHACSHGLVNSHLTSSFCRGLFSVLSTQS 740
Qy 456 MT-----DP-----
Db 741 LTELDSNLSLGDPGMRVLCETLQHPGNCIRRLWLGRGCLSHCECCFDISLVSSNQKIVE 800
Qy 460 -----LCHLSLTLSHCKLPDAVCRDLSEALRAAPALTEGL 496
Db 801 LLSNALGDFGIRLLCVGLKHLNCKLKLWLVSCLLTSACODLASVLSTSHSLRLYV 860
Qy 497 LHNRLSEAGRLMSEGLAMPQCRVQTVRV----- 525
Db 861 GENALGDSGVALICEKAKNPQCNLQKGLVNSGLTSVCCSALSSVLSTNQNLTHLYLGN 920
Qy 526 QLPDPORGLOYLIV-GMLROSPALTTLDLSGCCOLPAPM----- 561
Db 921 TLGD--KGIKLCEGILLHPDCKLQVLELDNCLNLTSHCCWDLSTLTSSQSLRKLSLGNND 978
Qy 562 -----VTYLCVAVLQHOGCGLOTLASVELSSEQLQAVKRAKPDVVI 606
Db 979 LGDLGWMFCEVLKQOSCLQNLGLSEMFTYETKSALETLOEKEPELTV 1028

RESULT 9
US-10-127-516-5
; Sequence 5, Application US/10127516
; Publication No. US20020187922A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Manji, Gulam A.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-370001
; CURRENT APPLICATION NUMBER: US/10/127,516
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 10/027,629
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 09/964,955
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506, 067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-127-516-5

Query Match 24.9%; Score 800.5; DB 14; Length 1034;
Best Local Similarity 28.4%; Pred. No. 5.1e-67;
Matches 219; Conservative 117; Mismatches 241; Indels 193; Gaps 17;

Qy 1 MLAQFORLLFIIDGADLPALGPPEAAP-CTOPFEAASGARVLGGLLSKALPTALLVLT 59
Db 288 IVKPSRILFLMDGFDLOGAFDEHIGPLCTDQKAERGDILLSSLRKKLPEASLLIT 347
Qy 60 TRAAAPGRLOGRLCSQCAEVGFSKDKKKYFYKFFRDRERAERAYRFVKNETLFLALC 119
Db 348 TRVALEKQLHLDHPRHVEILGFSSEAKKEYFFKYFSDQAQAAAFSLIOENEVLFTMC 407
Qy 120 FVPFVCWIVCTVLRQOLEGRDLRSRTKTTTSVYLLFITSVLSSAPVADGPRLOGLDRLN 179
Db 408 FIPLVCMIVCTGLKQMESGKSLAQTSKTTITVYVFFLSLLQPRGGSGEHLCAHLWGL 467
```

```
QY 180 CRLAREGVLCRRACFAEKLEFQLELRGSKVQTLFLSKKELPGVLTEVTYQIDFOSQEF 239
DB 468 CSLAADGIWQKILFEESDLRNLGQKADV-SAPLRNLFQKEVDEKFKYSFIHMTQEF 526
QY 240 LAALSLLDEGGVPRTAAG-----VGTLLRGAQ-PHSHLVLTTRFLGJLSAER 289
DB 527 FAMMYLLEEEKEGRNVGSRUKLPSRDVTVLLENYGKFGYLIFFVRFGLGVNQR 586
QY 290 MRDIERHFGCMVSRVKQBALRWVQGGCGPGVAPVTEGAKGLEDEEPEEESGEEP 349
DB 587 TSYLEKLSCKISQIRLELLKMI-----EVKAKAKKLO--IQPSQ----- 625
QY 350 NYPLELLYCLYETCEDAFVQALCRFPPELALORVFCRMDVAVLSYCVRCPCPAGCALRLI 409
DB 626 ---LELFYCLYEMQEEFVQAMDFPKIEIN--LSTRMDHMVSSFCIENCHRVSLSG 680
QY 410 SCRLVAAQKKKSLGKRLQ-----ASLGGSSQG--TTKQLPASLLHPLF-----QA 455
DB 681 FLHNPKEEEKEGRHLDMVQCVLPSSSHAACSHGLVNSH--TSFCRGLFSVLSTQS 740
QY 456 MT-----DP----- 459
DB 741 L*ELDLSDNSLGPGRMVLCTLOHPGCNIRRLWLGRCGLSHECCFDISLVSSNQKLVE 800
QY 460 -----LCHLSLTL*SHCKLPDAVCRDLSEALRAAPALTELGL 496
DB 801 LDLSNALGDFGIRLLCVGLKHLNKLWLVSCCLTSACCODLASVLSTSHSUTRLYV 860
QY 497 LHNRLSEAGRLMSEGLAWPQCRVQTVRV----- 525
DB 861 GENALGDSGVAILECEKAKNPOCNLOKGLVNSGLTSVCCSALS SVLSTNQNLTHLYLRGN 920
QY 526 QLPDPQRLQYLV-GMLRQSPALTTLDLSCGOLPAPM----- 561
DB 921 TLGD--KGIL*CEGLLHPDCKLVLELDNCNLTSHCCWDLSTLTSSQSLRKL*SGNND 978
QY 562 -----VTYLCVAVLHQCGGLQTLASVLSVSEQSLOELQAVKRAKPDIVI 606
DB 979 LGDLGVMMFCEVLKQOQSCLLQNLGLSEMYFN*YETKSALETLOEBKPELTV 1028

RESULT 11
US-10-027-629-5
; Sequence 5, Application US/10027629
; Publication No. US20020197660A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Manjil, Gulam A.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-34100
; CURRENT APPLICATION NUMBER: US/10/027,629
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 09/964,955
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506,067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-027-629-5
Query Match 24.9%; Score 800.5; DB 14; Length 1034;
Best Local Similarity 28.4%; Pred. No. 5,le-67;
Matches 219; Conservative 117; Mismatches 241; Indels 193; Gaps 17;
QY 1 MLAQQRLLIFLDGAELPALGPEAP-CTDPFAASGARVLGGJLSKALLPTALLIV* 59
```

```
DB 288 IVRPSRILFMDGDFDELOQAFDEHIGPLCTDWQKAERGDIILSSLRKKJLPPASLLIT 347
QY 60 TRAAPQRQLQRLCSPQCAEVRGFSKDKKKYFYKFRDERRAERAYRFRKENEITL*FALC 119
DB 348 TRVALEKLOHLLDHPHVLGFESEAKRKEYFKYFSDAQAPAA*SLIOENEVLTYC 407
QY 220 FVPVCMIVCTVLBOQLEGRDLSRTSKTTTSVLLTITSVLSAPVADGPRLOGDRLN 179
DB 408 FPLVCMIVCTGLKQOMESGKSLAQTSKTTTAVVVFVLSLLQPRGGSQEHGLCAH*WGL 467
QY 180 CRLAREGVLCRRACFAEKLEFQLELRGSKVQTLFLSKKELPGVLTEVTYQIDFOSQEF 239
DB 468 CSLAADGIWQKILFEESDLRNLGQKADV-SAPLRNLFQKEVDEKFKYSFIHMTQEF 526
QY 240 LAALSLLDEGGVPRTAAG-----VGTLLRGAQ-PHSHLVLTTRFLGJLSAER 289
DB 527 FAMMYLLEEEKEGRNVGSRUKLPSRDVTVLLENYGKFGYLIFFVRFGLGVNQR 586
QY 290 MRDIERHFGCMVSRVKQBALRWVQGGCGPGVAPVTEGAKGLEDEEPEEESGEEP 349
DB 587 TSYLEKLSCKISQIRLELLKMI-----EVKAKAKKLO--IQPSQ----- 625
QY 350 NYPLELLYCLYETCEDAFVQALCRFPPELALORVFCRMDVAVLSYCVRCPCPAGCALRLI 409
DB 626 ---LELFYCLYEMQEEFVQAMDFPKIEIN--LSTRMDHMVSSFCIENCHRVSLSG 680
QY 410 SCRLVAAQKKKSLGKRLQ-----ASLGGSSQG--TTKQLPASLLHPLF-----QA 455
DB 681 FLHNPKEEEKEGRHLDMVQCVLPSSSHAACSHGLVNSH--TSFCRGLFSVLSTQS 740
QY 456 MT-----DP----- 459
DB 741 L*ELDLSDNSLGPGRMVLCTLOHPGCNIRRLWLGRCGLSHECCFDISLVSSNQKLVE 800
QY 460 -----LCHLSLTL*SHCKLPDAVCRDLSEALRAAPALTELGL 496
DB 801 LDLSNALGDFGIRLLCVGLKHLNKLWLVSCCLTSACCODLASVLSTSHSUTRLYV 860
QY 497 LHNRLSEAGRLMSEGLAWPQCRVQTVRV----- 525
DB 861 GENALGDSGVAILECEKAKNPOCNLOKGLVNSGLTSVCCSALS SVLSTNQNLTHLYLRGN 920
QY 526 QLPDPQRLQYLV-GMLRQSPALTTLDLSCGOLPAPM----- 561
DB 921 TLGD--KGIL*CEGLLHPDCKLVLELDNCNLTSHCCWDLSTLTSSQSLRKL*SGNND 978
QY 562 -----VTYLCVAVLHQCGGLQTLASVLSVSEQSLOELQAVKRAKPDIVI 606
DB 979 LGDLGVMMFCEVLKQOQSCLLQNLGLSEMYFN*YETKSALETLOEBKPELTV 1028

RESULT 11
US-09-996-617-2
; Sequence 2, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1429
```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-931-071-2

Query Match 23.7%; Score 762; DB 10; Length 1429;
Best Local Similarity 33.2%; Pred. No. 4e-63;
Matches 220; Conservative 103; Mismatches 207; Indels 132; Gaps 22;

QY 1 MIAQPORLLFDGDEP--ALGPEAPCTDPEAASGARVLGGLLSKALLPTALLLV 58
DB ILRSBERLLFDLGVDE-PGWVLPSSSELCHWSQPADALLGSLGKTLIPASFLI 456
QY 59 TTRAAPGRLOGLRCSPOCAEVRGSDKKKYFKFRDERRAERAYRFVKNETLAL 118
DB TARTTALQNLIPSLQARWVEVLGFSSESRKEYFYRYFTDERQAIRAFRLVKSNEJMAL 516
QY 119 CFVPVCMIVCTVLRQQLLEGRDLRSRTSKTTSVLLFITSVLASAPVADGPRLOGLDRLN 178
DB CLVPWVSLACTCLMQMKRKEKLTLSKTTTTLCHVLAQALQAQPL--GPO---LRD 570
QY 179 LCRLAREGVLRRAQFAEKELEQLERLRSKVQTLFSLKKE--PGVLETE---VTQF--DOS 235
DB LCRLAREGVLRRAQFAEKELEQLERLRSKVQTLFSLKKE--PGVLETE---VTQF--DOS 235
QY 571 LCRLAREGVLRRAQFAEKELEQLERLRSKVQTLFSLKKE--PGVLETE---VTQF--DOS 235
DB LCRLAREGVLRRAQFAEKELEQLERLRSKVQTLFSLKKE--PGVLETE---VTQF--DOS 235
QY 236 FQEFALALSYLLEDDGVPRTAAGVGTLLRGDAQPHSHLVL-----KHSNCIIDLKTLKLEAYGIGHLFGAST 670
DB FQEFALALSYLLEDDGVPRTAAGVGTLLRGDAQPHSHLVL-----KHSNCIIDLKTLKLEAYGIGHLFGAST 670
QY 278 TRFLFGLLSAERMRIERHFGCMVSRVKQALRWVQGGCGPVAPEVTEGAKGLEDT 337
DB TRFLFGLLSAERMRIERHFGCMVSRVKQALRWVQGGCGPVAPEVTEGAKGLEDT 337
QY 671 TRFLFGLLSDEGEREMENIFHCLRSQ--GRNLMQWV-----PSLQLL 710
DB TRFLFGLLSDEGEREMENIFHCLRSQ--GRNLMQWV-----PSLQLL 710
QY 338 BEPEEEEEENYPLELLYCLYETQEDAFVQALCRPELALQVRFC---RMDVAVLS 394
DB BEPEEEEEENYPLELLYCLYETQEDAFVQALCRPELALQVRFC---RMDVAVLS 394
QY 711 LQP-----HSLSHCLYETRNKTFLOVMAHFEEMGX---CVETDVELLYCT 755
DB BEPEEEEEENYPLELLYCLYETQEDAFVQALCRPELALQVRFC---RMDVAVLS 394
QY 395 YVRCPCPAGQALRLISCR-----LVAAQKKKKSLGKRLQ 429
DB FCIKFSRHVKKLQLEGRHRSPTWVLFVRWVPTDAYWQILFSLVKVTRNL-KELD 814
QY 430 ASLGGSSGGTQKQIPASLLHPLFOAMTDPLCHLSSLTLSHCKLPDVCRLSEALRAAP 489
DB LS---GNS-----LSHSAVKSCLTLRRPCLLETJRLAGCGLTAEDCKDLAFGLRANQ 865
QY 490 ALTELGLLHNLSEAGLRMLSEGLAWPOCRVQTVRVLQDPORGL-----OYLVMGLROS 544
DB TLTEDLSFNVLTDAGAKHLQRLQPSCKLQ--RLQL--VSCGLTSDCCCLASVLSAS 921
QY 545 PALTTLDLSGCGLPAPMVTYLCVLIHQHOGCGLOTLASVLSVSELSQLOELQAVKRAKPD 604
DB PSLKELDLQNNLDVGVRLCEGLRHHPACKLIRLGLDQTLSDEMROELRALEQKXPOL 981
QY 605 VI 606
DB 982 LI 983

RESULT 12

US-09-931-071-2

; Sequence 2, Application US/09931071

; Patent No. US20020128219A1

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; APPLICANT: Alnemri, Enad S.

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED

; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF

; FILE REFERENCE: 07334-335001

; CURRENT APPLICATION NUMBER: US/05/931, 071

; CURRENT FILING DATE: 2002-03-18

; PRIOR FILING DATE: 1999-10-27

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-931-071-2

Query Match 23.7%; Score 762; DB 10; Length 1429;
Best Local Similarity 33.2%; Pred. No. 4e-63;
Matches 220; Conservative 103; Mismatches 207; Indels 132; Gaps 22;

QY 1 MIAQPORLLFDGDEP--ALGPEAPCTDPEAASGARVLGGLLSKALLPTALLLV 58
DB ILRSBERLLFDLGVDE-PGWVLPSSSELCHWSQPADALLGSLGKTLIPASFLI 456
QY 59 TTRAAPGRLOGLRCSPOCAEVRGSDKKKYFKFRDERRAERAYRFVKNETLAL 118
DB TARTTALQNLIPSLQARWVEVLGFSSESRKEYFYRYFTDERQAIRAFRLVKSNEJMAL 516
QY 119 CFVPVCMIVCTVLRQQLLEGRDLRSRTSKTTSVLLFITSVLASAPVADGPRLOGLDRLN 178
DB CLVPWVSLACTCLMQMKRKEKLTLSKTTTTLCHVLAQALQAQPL--GPO---LRD 570
QY 179 LCRLAREGVLRRAQFAEKELEQLERLRSKVQTLFSLKKE--PGVLETE---VTQF--DOS 235
DB LCRLAREGVLRRAQFAEKELEQLERLRSKVQTLFSLKKE--PGVLETE---VTQF--DOS 235
QY 571 LCRLAREGVLRRAQFAEKELEQLERLRSKVQTLFSLKKE--PGVLETE---VTQF--DOS 235
DB LCRLAREGVLRRAQFAEKELEQLERLRSKVQTLFSLKKE--PGVLETE---VTQF--DOS 235
QY 236 FQEFALALSYLLEDDGVPRTAAGVGTLLRGDAQPHSHLVL-----KHSNCIIDLKTLKLEAYGIGHLFGAST 670
DB FQEFALALSYLLEDDGVPRTAAGVGTLLRGDAQPHSHLVL-----KHSNCIIDLKTLKLEAYGIGHLFGAST 670
QY 278 TRFLFGLLSAERMRIERHFGCMVSRVKQALRWVQGGCGPVAPEVTEGAKGLEDT 337
DB TRFLFGLLSAERMRIERHFGCMVSRVKQALRWVQGGCGPVAPEVTEGAKGLEDT 337
QY 671 TRFLFGLLSDEGEREMENIFHCLRSQ--GRNLMQWV-----PSLQLL 710
DB TRFLFGLLSDEGEREMENIFHCLRSQ--GRNLMQWV-----PSLQLL 710
QY 338 BEPEEEEEENYPLELLYCLYETQEDAFVQALCRPELALQVRFC---RMDVAVLS 394
DB BEPEEEEEENYPLELLYCLYETQEDAFVQALCRPELALQVRFC---RMDVAVLS 394
QY 711 LQP-----HSLSHCLYETRNKTFLOVMAHFEEMGX---CVETDVELLYCT 755
DB BEPEEEEEENYPLELLYCLYETQEDAFVQALCRPELALQVRFC---RMDVAVLS 394
QY 395 YVRCPCPAGQALRLISCR-----LVAAQKKKKSLGKRLQ 429
DB FCIKFSRHVKKLQLEGRHRSPTWVLFVRWVPTDAYWQILFSLVKVTRNL-KELD 814
QY 430 ASLGGSSGGTQKQIPASLLHPLFOAMTDPLCHLSSLTLSHCKLPDVCRLSEALRAAP 489
DB LS---GNS-----LSHSAVKSCLTLRRPCLLETJRLAGCGLTAEDCKDLAFGLRANQ 865
QY 490 ALTELGLLHNLSEAGLRMLSEGLAWPOCRVQTVRVLQDPORGL-----OYLVMGLROS 544
DB TLTEDLSFNVLTDAGAKHLQRLQPSCKLQ--RLQL--VSCGLTSDCCCLASVLSAS 921
QY 545 PALTTLDLSGCGLPAPMVTYLCVLIHQHOGCGLOTLASVLSVSELSQLOELQAVKRAKPD 604
DB PSLKELDLQNNLDVGVRLCEGLRHHPACKLIRLGLDQTLSDEMROELRALEQKXPOL 981
QY 605 VI 606
DB 982 LI 983

RESULT 13

US-10-028-374-15

; Sequence 15, Application US/10028374

; Publication No. US20030143706A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED

; TITLE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLREMI

; FILE REFERENCE: D0067NP

; CURRENT APPLICATION NUMBER: US/10/028,374

; CURRENT FILING DATE: 2001-12-20

; PRIOR FILING DATE: 2000-12-22

; NUMBER OF SEQ ID NOS: 22

; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/259,479
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US 60/260,616
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-392-11

Query Match 23.7%; Score 762; DB 15; Length 1429;
 Best Local Similarity 33.2%; Pred. No. 4e-63;
 Matches 220; Conservative 103; Mismatches 207; Indels 132; Gaps 22;

QY	1	MLAOPORLLFILDGADLP--ALGGPEAPCTDPFEAASGARVLGGLLSKALLPTAALLV	58
DB	398	LSRPERLLFILDGVDE-PGWLQEPSSSELCLHWSQPQPAQALLGSLGKTLPEASFL	456
QY	59	ITRAAAPGRLOGLCSQCAEVRGFSDDKXKYFKFRDERRAERAYRFVKENETL	118
DB	457	TARTTALQNLIPSLQARWVEVLGFSESRKEYFYFYFDCQAIRAPFLVKSRELWAL	516
QY	119	CFVPFCWIVCTVLRQOLEGRDLRSRTSTTSVLLFITSVLSSAPVADGPRLQGDLEN	178
DB	517	CLVPVWSWLACTCLWQMKKXELTSTTTTLCYHVAQALQAOPL--GPG	570
QY	179	LCRLAREGVLGRRAPAEKEJOLELRGSKVOTFLSKKELPGVLETE---VTYOFIDQS	235
DB	571	LCSLAAEGTWOKTLFSPDRLKKGIDGAIIST-FLKM---GILQEHPIPLSYSFHLC	625
QY	236	FOFLAANLYLLEDDGVPTAGGVGTLRLGCAQPHSHLVL-----T	277
DB	626	FOEFAAMSIVLEDD-----EKGRG-----KESNCIIDLEKTLSEAYGIHGLFGAST	670
QY	278	TRFLGLLSAERMRIERHFGGMVSERVKQEAALRWVQGGCGPGVAPEVTEGAKGLFDT	337
DB	671	TRFLGLLSDEGEREYENIFHCLSQ--GRNLMQWV-----PSLQLL	710
QY	338	EEPEEEEGEENYPLELLYC-YETCEDAFVQALCRFPALQVRFC---RMDVAVLS	394
DB	711	LQP-----HSLDSLHC-YETRNKTFELCVMAHPEEMGM-----CVETDMELLVCT	755
QY	395	YCVRCPCPAGQALRLISCR-----LVAAQEKKKKS-LGKELQ	429
DB	756	FCIKFSRHVKLQLLEGROHRTWPTMVVLFRWVFTDAYWQILFVLKVTNRN-KEJD	814
QY	430	ASLGGGSSQGTTKQLPASLLHPLFQAMTDPLCHLSLTL-SHCKLDPDAVCROLSEALRAAP	489
DB	815	LS-----LSHSAVKSCLKTLRRPRCLLETRLAGCGLTAEDCKDLAFGLRANQ	865
QY	490	ALTELGLLNRLSEAGRLM-SEGLAWPQCRVQTVRVQLPDQORGL-----QYLVGMLRQS	544
DB	866	TLTLDLSFNVLTDAGAKLQORLQPSCKLQ--KQL--VSCGLTSDCCQCLASVLSAS	921
QY	545	PALTTLDLSCCOLPAPMTVYLCVAVLQHCQCGQLQTLASVEVSELOSLOAVKRAKFDL	604
DB	922	PSLKLDELQONNDDVGVRLLCEGLRHPACKLRGLDOTTLSDEMRELRALEGEKPOL	981
QY	605	VI	606
DB	982	LI	983

Search completed: October 2, 2003, 18:00:42
Job time : 75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - nucleic search, using frame_plus_p2n model

Run on: October 3, 2003, 15:49:16 ; Search time 410 Seconds
(without alignments)
3888.511 Million cell updates/sec

Title: US-10-029-347-2

Perfect score: 3218

Sequence: 1 MIAQPORLLFILDGADELPA.....ITHPALDGHPPKELISTP 625

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1708419 seqs, 1275431651 residues

Total number of hits satisfying chosen parameters: 3416838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_6/ptodata/1/pubpna/us07_PUBCOMB.seq
-DB=PublishedApplications_NA -GPM=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OJTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0
-NAXLEN=200000000 -USER=US10029347 -CGCN 1 1 221 -runat 02102003 173438 17146
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:

1: /cgn2_6/ptodata/1/pubpna/us07_PUBCOMB.seq
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3098	96.3	2844	11	US-09-965-621-17 Sequence 17, Appl

2	926.5	25.7	3108	11	US-09-965-621-23	Sequence 23, Appl
3	825.5	25.7	3186	14	US-10-066-521-17	Sequence 17, Appl
4	800.5	24.9	3102	12	US-10-132-967-6	Sequence 6, Appl
5	800.5	24.9	3102	13	US-10-127-516-6	Sequence 6, Appl
6	800.5	24.9	3102	13	US-10-027-629-6	Sequence 6, Appl
7	800.5	24.9	3857	12	US-10-132-967-4	Sequence 4, Appl
8	800.5	24.9	3857	13	US-10-127-516-4	Sequence 4, Appl
9	800.5	24.9	3857	13	US-10-027-629-4	Sequence 4, Appl
10	762	23.7	4287	10	US-09-996-617-5	Sequence 5, Appl
11	762	23.7	4287	10	US-09-931-071-5	Sequence 5, Appl
12	762	23.7	4422	10	US-09-388-221-1	Sequence 1, Appl
13	762	23.7	4556	10	US-09-388-221-9	Sequence 9, Appl
14	762	23.7	5444	10	US-09-996-617-1	Sequence 1, Appl
15	762	23.7	5444	10	US-09-931-071-1	Sequence 1, Appl
16	762	23.7	5444	11	US-09-956-712-3	Sequence 3, Appl
17	762	23.7	5531	11	US-09-956-712-11	Sequence 11, Appl
18	760	23.6	5100	11	US-09-956-712-12	Sequence 12, Appl
19	716	22.2	4200	10	US-09-388-221-3	Sequence 3, Appl
20	716	22.2	4332	10	US-09-388-221-5	Sequence 5, Appl
21	716	22.2	4466	10	US-09-388-221-11	Sequence 11, Appl
22	694	21.6	4931	12	US-10-028-374-1	Sequence 1, Appl
23	694	21.6	4931	12	US-10-183-770-1	Sequence 1, Appl
24	691.5	21.5	2524	11	US-09-965-621-58	Sequence 58, Appl
25	650	20.2	2940	9	US-09-848-035-14	Sequence 14, Appl
26	650	20.2	2940	10	US-09-986-224-14	Sequence 14, Appl
27	650	20.2	3263	9	US-09-848-035-12	Sequence 12, Appl
28	650	20.2	3263	10	US-09-986-224-12	Sequence 12, Appl
29	634	19.7	3150	12	US-10-239-663-9	Sequence 9, Appl
30	619.5	19.3	3099	12	US-10-132-967-3	Sequence 3, Appl
31	619.5	19.3	3099	13	US-10-127-516-3	Sequence 3, Appl
32	619.5	19.3	3099	13	US-10-027-629-3	Sequence 3, Appl
33	619.5	19.3	3431	12	US-10-132-967-1	Sequence 1, Appl
34	619.5	19.3	3431	13	US-10-127-516-1	Sequence 1, Appl
35	619.5	19.3	3431	13	US-10-027-629-1	Sequence 1, Appl
36	616.5	19.2	3189	12	US-10-239-663-10	Sequence 10, Appl
37	598.5	18.6	2575	9	US-09-848-035-7	Sequence 7, Appl
38	598.5	18.6	2575	10	US-09-986-224-7	Sequence 7, Appl
39	598.5	18.6	2982	14	US-10-066-521-25	Sequence 25, Appl
40	598.5	18.6	2985	11	US-09-965-621-15	Sequence 15, Appl
41	598.5	18.6	3368	14	US-10-066-521-23	Sequence 23, Appl
42	597.5	18.6	3048	10	US-09-986-224-20	Sequence 20, Appl
43	597.5	18.6	3051	10	US-09-986-224-18	Sequence 18, Appl
44	586.5	18.2	2691	14	US-10-066-521-21	Sequence 21, Appl
45	566	17.6	96649	11	US-09-956-712-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-965-621-17
; Sequence 17, Application US/09965621
; Publication No. US20030077699A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/09/965,621
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 2844
; TYPE: DNA
; ORGANISM: Homo sapien

FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)...(2841)
 US-09-965-621-17

Alignment Scores:

Prod. No.: 0 Length: 2844
 Score: 3098.00 Matches: 613
 Percent Similarity: 90.16% Conservative: 1
 Best Local Similarity: 90.01% Mismatches: 11
 Query Match: 96.27% Indels: 56
 DB: 11 Gaps: 2

US-10-029-347-2 (1-625) x US-09-965-621-17 (1-2844):

QY : MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeuProAla 20
 DB 799 ATGCTGGCCGAGCGGAGCGGCTGCTCTCATCTGGACGGCGGAGCGAGCTGGCGCG 858
 QY 21 LeuGlyGlyProGluAlaProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
 DB 859 CTGGGGGGCCCGAGGCGCGCCCTGACAGACCCCTTCAGCGCGCGAGCGCGCGCG 918
 QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
 DB 919 GTGCTAGGCGGCTGCTGAGCAAGCGCTGCTGCCACGGCCCTCTGCTGGTACCACG 978
 QY 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
 DB 979 CGCGCGCGCGCCCGGAGGCTGCGAGCGCGCGCTGTGTCCCGCAGTGGCGCGAGG 1038
 QY 81 ArgGlyPheSerAspLysAspLysLysLysLysLysPheThrLysPhePheArgAspGluArg 100
 DB 1039 CGCGGCTTCTCCGACAAAGGCAAGAAAGATATTTCTACAGTTCTTCGGGATAGAGG 1096
 QY :01 ArgAlaGluArgAlaThrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
 DB 1099 AGGCGCGAGCGCGCTACCGTTCGTGAGGAGCAAGCAGCGCTGTTCGGCTGTGCTTC 1158
 QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuLeuGlyArg 140
 DB 1159 GTGCCCTTCGTGTGCTGGATGCTGTGTCACCGTGTGCGCAGCAGCTGGAGCTCGTGG 1218
 QY 141 AspLeuSerArgThrSerLysThrThrSerValThrLeuLeuPheIleThrSerVal 160
 DB :219 GACCTGTGCGGCGCTGTCACAGACCAACACGTCAGTGTACTCTCTTTCATCACCGCGT 1278
 QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlnGlyAspLeuArgAsnLeuCys 180
 DB :279 CTGAGCTCGCTCCGCTAGCGGACGGCGCGCGCTGTCAGGGCGACCTGCGCAATCTGTGC 1338
 QY 181 ArgLeuAlaArgGluGlyValIleGlyArgArgAlaGlnPheAlaGlyLeuGlu 200
 DB 1339 CGCTGGCCCGCGAGGCGCTCTCGAGCGGCGGCGCAGTTTTCGCGAGAGGAACCTGGAG 1398
 QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
 DB 1399 CAACCTGGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTCTCAGCAAAAAGAGGCTGCCG 1458
 QY 221 GlyValLeuGluThrGluValThrThrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
 DB 1459 GCGGTGCTGGAGACAGAGTTCACCTACCACTTTCATCGACCGAGCTTCGCGAGAGGTTCTTC 1518
 QY 241 AlaAlaLeuSerThrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
 DB :519 SCGGCACTGTCTTACCTGCTGGAGGACGGCGGGTGCACGAGACCGCGCTGCGCGCGTT 1578
 QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
 DB 1579 GGGACACTCTCTGGTGGGAGCGCCAGCGCACAGGCACTTGTGCTCACCGCGCTTC 1638
 QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300

DB 1639 CTCCTCGGACTGCTGAGCGCGGAGCGGATGCGGCACATCGAGCGCCACTTCGGCTGCATG 1698
 QY 301 ValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlnGlyGlnGlyCys 320
 DB 1699 GTTTCAGAGCGTGTGAAGCAGGAGCGCCCTCGCGTGGGTGCGAGGACAGGACAGGCGCTG 1758
 QY 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
 DB 1759 CCCGAGTGGCACAGAGGTGACCCGAGGGGCGCAAGGGCTTCGAGGACACCGAAGAGCCA 1818
 QY 341 GluGluGluGluGluGlyGlyGluGluProAsnThrProLeuGluLeuLeuLeuLeuLeu 360
 DB 1819 GAG 1878
 QY 361 GluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
 DB 1879 GAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1938
 QY 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerThrCysValArgCysCys 400
 DB 1939 CAGCGAGTGGCTTCTGCGCATGGACGTGGCTGTCTGAGCTACTCGGTGAGGTGCTGC 1998
 QY 401 ProAlaGlyGlnAlaLeuArgLeuIleSerCysArgLeuValAlaAlaGlnGluLys 420
 DB 1999 CCTGCTGGACAGGCACTGGCGCTGATCAGCTGCAGATTGGTTGCTGCGCAGGAGAGAAG 2058
 QY 421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlySerSerGlnGlyThr 440
 DB 2059 AAGAAGAGCTTGGGAAAGCGGCTCCAGGCGAGCTGGGTGGCGCAGCTGGCTGGGGACC 2118
 QY 440 ----- 440
 DB 2119 CAACCTGGCTCCAGAGTACCTTTTCGACCACCTCTGTCGACATCTGCCCCACACTCCA 2178
 QY 440 ----- 440
 DB 2179 CCAGACCTCGGCTCTCCAGGCAAGGCTTTTGCCAGAGTTCTCTTTGAATATAGCTCCA 2238
 QY 441 ThrLysGlnLeuProAlaSerLeuLeu----- 449
 DB 2239 ATTCAGCCCTCGCCAGGGCTTGGCATCTGTTGAGAGATGAATGTACGGTGTGGCA 2298
 QY 450 -----HisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSer 464
 DB 2299 GGGCTGGCGCTGGGGACCCCAAGACCATGACTGACCCACTGTGCACTCTGAGC 2358
 QY 465 SerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerGluAla 484
 DB 2359 AGCCTCAGCTGTCCCACTGCAAACTCCCTGACCGGCTCTGCCGAGACCTTTCTGAGGCC 2418
 QY 485 LeuArgAlaAlaProAlaLeuThrGluLeuLeuHisAsnArgLeuSerGluAla 504
 DB 2419 CTGAGGCGAGCCCGGCACTGACGAGCTGGGCTCTCTCCACACAGGCTCAGTGGGCA 2478
 QY 505 GlyLeuArgMetLeuSerGluGlyLeuAlaThrProGlnCysArgValGlnThrValArg 524
 DB 2479 GGACTCGGTATGCTGAGTGAGGCGCTAGCTGGCGCGCAGTGCAGGGTGCAGCGGTGAGG 2538
 QY 525 ValGlnLeuProAspProGlnArgGlyLeuGlnThrLeuValGlyMetLeuArgGlnSer 544
 DB 2539 GTACAGCTGCTGACCCCGGAGGCTCCAGTCTGTTGGGTATGCTTTCGCGCAGAGC 2598
 QY 545 ProAlaLeuThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrThr 564
 DB 2599 CTTGCCCTGACCACTTGGATCTCAGCGGTGCCACTGCGCGCCCGCCCTGCTGACCTAC 2658
 QY 565 LeuCysAlaValLeuGlnHisGlnGlyLeuGlnThrLeuSerLeuAlaSerVal 584
 DB 2659 CTGTGTGAGTCTCTGACGACCAAGGATGGCGCTGCAGACCTCAGTGTGGCTCTGTG 2718
 QY 585 GlnLeuSerGluGlnSerLeuGlnGluGlnAlaValLysArgAlaLysProAspLeu 604
 DB 2719 GAGCTGAGCGAGCAGTCACTACAGGAGCTTCAGGCTGTGAAGAGAGCAAGCGGATCTG 2778

```

QY 605 ValIleThrHisProAlaLeuAspGlyHisProGlnProProIysGluLeuIleSerThr 624
DB 2779 GTATACACACCCAGCGCTGAGCCACCCACCAACCTCCCAAGAACTCATCTCGACC 2838
QY 625 Phe 625
DB 2839 TTC 2841

RESULT 2
US-09-965-621-23
; Sequence 23, Application US/09965621
; Publication No. US20030077699A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 3108
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3105)
US-09-965-621-23

Alignment Scores:
Pred. No.: 3,46e-84 Length: 3108
Score: 826.50 Matches: 231
Percent Similarity: 48.00% Conservative: 105
Best Local Similarity: 33.00% Mismatches: 247
Query Match: 25.68% Indels: 1-9
DB: 17

US-10-029-347-2 (1-625) x US-09-965-621-23 (1-3108)
QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeu---Pro 19
DB 763 CTCATCGAGTTCCCGAGCGCCTCTTTTCATCATCGAGCGCTTCGAGCTCAAGCT 822
QY 20 AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaIleSerGlyAla 39
DB 823 TCTTTCACGATCCTCAGGACCTCGTGCCTCTGCTGGAGGAGAAACGGCCACGGAG 882
QY 40 ArgValLeuGlyGlyLeuLeuSerLysAlaLeuProThrAlaLeuLeuValThr 59
DB 883 CTGTTCTTAACAGCTTAATTCGAGAGAGCTCTCCCTGAGCTATCTTTGCTCATCACC 942
QY 60 ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu 79
DB 943 ACACGGCCACGGCTTTGGAGAGCTCCACCGTCTGCTGGAGCACCCAGGCATGTGGAG 1002
QY 80 ValArgGlyPheSerAspLysLysLeuLysTyPheTyLysPhePheArgAspGlu 99
DB 1003 ATCTCGGGCTCTCTCAGGAGAGAGAGAGAAATCTTCTACAAGTATTTCCCAATGCA 1062
QY 100 ArgArgAlaGluArgAlaTyArgPheValLysGluAsnGlnThrLeuPheAlaLeuCys 119
DB 1063 GAGCAGCGGGCCNAGTCTTCAATTACGTGAGGAGCAACAGCGCTCTCTTACCACTGTGC 1122

```

```

QY 120 PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGly 139
DB 1123 TTGTCCTCCCTGTGTGCTGGG:GGTGT:ACTGCTCCAGCAGCAGCTGGAGGGTGG 1182
QY 140 ArgAspLeuSerArgThrSerLysThrThrThrSerValThrLeuPheIleThrSer 159
DB 1183 GGGCTGTTGAGCAGACGCTCCAGGACCACCTGCAGTGTATCATCTTACCTGCTGAGT 1242
QY 160 ValLeuSerSerAlaProValAlaAspGlyProArgLeuGln-----GlyAspLeuArg 177
DB 1243 CTGATCAACCCCAAGCGGGGGCC-----CGCGCCTCCAGCCCCCACCACCAAGAGA 1296
QY 178 AsnLeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLys 197
DB 1297 GGGTTGTCTCTTGGCGCAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAG 1356
QY 198 GluLeuGluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLys 217
DB 1357 GACCTCGGAAGCAGCGCTAGACGGGAAGAGCTC---TCTGCCTTCTCAACATGAAC 1413
QY 218 GluLeuProGlyValLeuGluThrGluValThrTyGlnPheIleAspGlnSerPheGln 237
DB 1414 ATCTTCAGAGGACATCAACTGTGTGAGAGGTACTACAGCTTTCATCCACTTGAGTTTCCAG 1473
QY 238 GluPheLeuAlaAlaLeuSerTyLeuLeuGluAspGlyGlyValProArgThrAlaAla 257
DB 1474 GAATTTCTTGCAGCTATGTACTATATCTCTGGAGAGGG------GAG 1515
QY 258 GlyGlyValGly-----ThrLeuLeuArgGlyAspAlaGlnPro 270
DB 1516 GCGGGGCGAGCCAGCAGCAGGAGCTGTGACCGCTGTTCACCGAGTACCGTTTCTGAA 1575
QY 271 HisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMet 290
DB 1576 AGGAGCTTCTCGCACTCACCAGCCCTCTCTGTTGGACTCTGACAGGAGACCCAGG 1635
QY 291 ArgAspIleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAlaLeu 310
DB 1636 AGCCACCTGGAGAAGAGTCTCTCTGGAAGTCTCGCCGACATCAAGATGGACCTGTTG 1695
QY 311 ArgTrpValGlnGlnGlnGlnGlnGlyCysProGlyValAlaProGluValThrGluGly 330
DB 1696 CAGTGGATCCAAAGCAAGCTCAG----- 1719
QY 331 AlaLysGlyLeuGluAspThrGluGluProGluGluGluGluGluGluGluProAsn 350
DB 1720 -----ACGAGCGCTCCACCTCGCAGCAGGCTCC----- 1749
QY 351 TyrProLeuGluLeuLeuTyCysLeuTyGluThrGlnGluAspAlaPheValArgGln 370
DB 1750 -----TTGAGTTCTTTCAGCTGCTGTAGAGATCCAGGAGGAGGAGTTTATCCAGCAG 1803
QY 371 AlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPheCysArgMetAspVal 390
DB 1804 GCCCTGAGCCACTTCCAGGTGATCGTGTGTCAGCAACATTGCC---TCCAGATGGAGCAC 1860
QY 391 AlaValLeuSerTyCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuIleSer 410
DB 1961 ATGCTCTCTCTGTTCTGTCTGAGCGCTCGAGGAGCGCCAGGTGCTGCACCTGTATGSC 1920
QY 411 CysArg----- 412
DB 1921 GCCACCTACAGCGCGGAGGAGAGCCGCGAGGTGCTCGCAGGAGCGCACAGCTG 1980
QY 413 LeuValAlaAlaGlnGluLysLysLysSerLeuGlyLysArgLeuGlnAlaSerLeu 432
DB 1981 TTGCT-CAAGCTCAGACAGAGAGGAGCGTTCGTCTGGA---CGCTACAGTGAACATCT 2036
QY 433 GlyGlyGly-----SerSerGlnGlyThrThrLysGlnLeuProAla-SerLeu 448
DB 2037 GGCAGCGGCTGTGACCAATCCAAACCTGTAGAGGTGCTCTCTGCTACCGAAATGCCCT 2096
QY 448 u-----LeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSe 464

```

```
Db 2097 GGCAGCGGGGTGAAGCTCTCTCAAGACTCAGACACCCCAACTGCACACTTCA 2156
Qy 464 rSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerGluAl 484
Db 2157 GAACCTGAGGCTGAAGAGGTSCCGCATCTCCAGCTCAGCGCTGGCAGGACCTCTCTCAGC 2216
Qy 484 aLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAspArgLeuSerGluAl 504
Db 2217 TCTCATAGCCAAATAGAAATTTGACAGGATGATCTCAGTGGCAAGCGGTGGATCC 2276
Qy 504 aGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrValAr 524
Db 2277 AGCATGATGCTGTTTGGCAGGGCCCTGGGCATCTCCCACTGAGGCTGCAGATGATTC 2336
Qy 524 gValGlnLeuProAspProGlnArgGlyLeuValGlyMetLeuArgG 543
Db 2337 GTTGAGGAAGTGCAGCTGAGTCCGGGGCTTTCAGGAGATGCTTCTGTGCTCGGCAC 2396
Qy 543 nSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValTr 563
Db 2397 CAACCCACATCTGGTTGAGTTGGACCTGACAGGAAATGCACATTTGGCGCTGAG 2454
Qy 563 rTyLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeu 579
Db 2457 GTTACTATGCCAGGGACTGAGGACCCAGCTCTGCAGACTACGACTTTGTGGCTGAAGAT 2516
Qy 580 -----SerLeuAlaSerValGluLeuSerCysGlnSe 590
Db 2517 CTCCCGCCTCAGCTGCTGCTGCTGTGAGAGCTGGGCTCAACTCTCAGTGTGAACAGAG 2576
Qy 590 rLeuGlnGluLeuGlnAlaValLysArgAlaLysProAspLeuValLe 606
Db 2577 CTTGAGAGACTGGACCTGAGCTGATGAGCTGGGGGACCTCGGGTGTCTGTGCTGTG 2636
Qy 607 -----ThrHisPro 609
Db 2637 TGAGGGCTCAGGCATCCACAGTGCAGAGCTCCAGACCTCGGTTGGGCATCTCGCGCT 2696
Qy 610 -----AlaLeuAspGlyHisProGlnProPro 618
Db 2697 GGGCTGTGCCCTGTGAGGCTTTCTGTGGTGTCTCA-GGCCAACCAACCTCCG 2753

RESULT 3
US-10-066-521-17
; Sequence 17, Application US/10066521
; Publication No. US20030027757A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PVRIN/NBS/LRR
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-334301
; CURRENT APPLICATION NUMBER: US/10/066,521
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/318,645
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/265,231
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(3183)
US-10-066-521-17
```

Alignment Scores:

Pred. No.: 4.68e-84 Length: 3186

```
Score: 825.50 Matches: 230
Percent Similarity: 48.07% Conservative: 106
Best Local Similarity: 32.90% Mismatches: 247
Query Match: 25.65% Indels: 118
DB: 14 Gaps: 17

US-10-029-347-2 (1-625) x US-10-066-521-17 (1-3186)

Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeu---Pro 19
Db 844 CTCATCCGAGTTCGCGAGCGCTCTTTTCATCATPCGACGGCTTCGATGAGCTCAAGCCT 903
Qy 20 AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAla 39
Db 904 TCTTTCCAGCATCTCTCAGGAGCCCTGTGCTCTGCTGGGAGGAGAAACGCCCCAGGAG 963
Qy 40 ArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 59
Db 964 CTGCTTCTTAACAGCTTAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCAC 1023
Qy 60 ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu 79
Db 1024 ACACGCCCCAGCGCTTTGGAGAGCTCCACCGTCTGCTGGAGACCCCGAGCACTGGGAG 1083
Qy 80 ValArgGlyPheSerAspLysAspLysLysTyPheTyLysPhePheArgAspGlu 99
Db 1084 ATCCTGGGCTTCTCTGAGCGCAGAAAGGAAGGAATATCTTACAAGTATTTCCACAATGCA 1143
Qy 100 ArgArgAlaGluArgAlaTyArgPheValLysGluAsnGluThrLeuPheAlaLeuCys 119
Db 1144 GAGCAGCGCGGCGCAAGTCTTCAATTACGTGAGGACACACGAGCGCTCTCTCACCATGTGC 1203
Qy 120 PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuLeuGly 139
Db 1204 TTCGTCCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1263
Qy 140 ArgAspLeuSerArgThrSerLysThrThrThrSerValTyLeuLeuPheIleThrSer 159
Db 1264 GGGCTGTTTGAGACAGACGCTCCAGGACCAACCTGCGAGTGTATCATGCTCTACCTGCTGAGT 1323
Qy 160 ValLeuSerSerAlaProValAlaAspGlyProArgLeuGln-----GlyAspLeuArg 177
Db 1324 CTGATGCAACCAAGCGCGGGGCGCC-----CCGGCGCTCCAGCGCCCGCCACCAACAGAGA 1377
Qy 178 AsnLeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLys 197
Db 1378 GGGTGTGCTCTCTGCGCGCAGATGGGCTCTGGAAATCAGAAATCTTATTTGAGGAGCAG 1437
Qy 198 GluLeuGluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLys 217
Db 1438 GACCTCCGGAAGCAGCGGCTAGACGGGGAAGACGCTC---TCTGCTTCTCTCAACATGAAC 1494
Qy 218 GluLeuProGlyValLeuGluThrGluValThrTyGlnPheIleAspGlnSerPheGln 237
Db 1495 ATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTGAGTTTCAG 1554
Qy 238 GluPheLeuAlaAlaLeuSerTyLeuLeuGluAspGlyGlyValProArgThrAlaAla 257
Db 1555 GAATCTTTGAGCTATGTACTATATCTCTGACGAGGGG-----GAG 1596
Qy 258 GlyGlyValGly-----ThrLeuLeuArgGlyAspAlaGlnPro 270
Db 1597 GSCGGGCGAGGCCAGACCGAGCGTGACCGGCTGTTGACCGAGTACGCGCTTTTCTCAA 1656
Qy 271 HisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMet 290
Db 1657 AGGAGCTTCTCTGGCACTCACCAGCGCTTCTGTTGAGACTCTCTGAACGAGGAGACCAAG 1716
Qy 291 ArgAspIleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAlaLeu 310
Db 1717 AGCCACTGGAGAGAGGTCTCTGTGGAAAGGTCTCGCGGCACATCAAGATGAGCTGTG 1776

311 ArgTrpValGlnGlnGlyGlnGlyCysProGlyValAlaProGluValThrGluGly 330
```

[illegible]

Dd 2778 CTCTGCCGCTGTGAGGGTCTTTTCGTGGTGCTCCA-GGCCAACACCAACCTCCG 2831

RESULT 4
US-10-132-967-6
; Sequence 6, Application US/10132967
; Publication No. US20030170841A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN PROTEIN FAMILY AND
; FILE REFERENCE: 07334-311001
; CURRENT APPLICATION NUMBER: US/10/132.967
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US/09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506,067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-967-6

Alignment Scores:

Pred. No.:	3,35e-81	Length:	3102
Score:	800.50	Matches:	219
Percent Similarity:	43.64%	Conservative:	117
Best Local Similarity:	28.44%	Mismatches:	241
Query Match:	24.88%	Indels:	193
DB:	12	Gaps:	17

US-1C-029-347-2 (1-625) x US-10-132-967-6 (1-3102)

Qy	1	MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuPcAla	20
Dd <td>862<td>ATCGTGGAAGAACCCTCCAGAATCCTCTTCTCATGCAGCGCTTCGATGAGCTGCAAGGT<td>921</td></td></td>	862 <td>ATCGTGGAAGAACCCTCCAGAATCCTCTTCTCATGCAGCGCTTCGATGAGCTGCAAGGT<td>921</td></td>	ATCGTGGAAGAACCCTCCAGAATCCTCTTCTCATGCAGCGCTTCGATGAGCTGCAAGGT <td>921</td>	921
Qy <td>21<td>LeuGlyGlyProGluAlaAlaPro---CysThrAspProPheGluAlaAlaSerGlyAla<td>39</td></td></td>	21 <td>LeuGlyGlyProGluAlaAlaPro---CysThrAspProPheGluAlaAlaSerGlyAla<td>39</td></td>	LeuGlyGlyProGluAlaAlaPro---CysThrAspProPheGluAlaAlaSerGlyAla <td>39</td>	39
Dd <td>922<td>GCCTTTGACGAGCACATAGGACCGCTCTGCACCTACTGGCAGAAAGCCGCGGGAGAC<td>981</td></td></td>	922 <td>GCCTTTGACGAGCACATAGGACCGCTCTGCACCTACTGGCAGAAAGCCGCGGGAGAC<td>981</td></td>	GCCTTTGACGAGCACATAGGACCGCTCTGCACCTACTGGCAGAAAGCCGCGGGAGAC <td>981</td>	981
Qy <td>40<td>ArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuLeuValThr<td>59</td></td></td>	40 <td>ArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuLeuValThr<td>59</td></td>	ArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuLeuValThr <td>59</td>	59
Dd <td>982<td>ATTCTCTTGACGAGCCTCATAGAAAAGAGTGCTTCCCAGAGCCTCTCTGCTCATCACC<td>104</td></td></td>	982 <td>ATTCTCTTGACGAGCCTCATAGAAAAGAGTGCTTCCCAGAGCCTCTCTGCTCATCACC<td>104</td></td>	ATTCTCTTGACGAGCCTCATAGAAAAGAGTGCTTCCCAGAGCCTCTCTGCTCATCACC <td>104</td>	104
Qy <td>60<td>ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu<td>79</td></td></td>	60 <td>ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu<td>79</td></td>	ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu <td>79</td>	79
Dd <td>1042<td>ACGAGACCTGTGGCCCTGGAGAACTGCACACTTGTCTGGNCCATCTCTGGCATGTGGAG<td>110</td></td></td>	1042 <td>ACGAGACCTGTGGCCCTGGAGAACTGCACACTTGTCTGGNCCATCTCTGGCATGTGGAG<td>110</td></td>	ACGAGACCTGTGGCCCTGGAGAACTGCACACTTGTCTGGNCCATCTCTGGCATGTGGAG <td>110</td>	110
Qy <td>90<td>ValArgGlyPheSerAspLysAspLysLysLysTyrrPheTyrllysPhePheArgAspGlu<td>99</td></td></td>	90 <td>ValArgGlyPheSerAspLysAspLysLysLysTyrrPheTyrllysPhePheArgAspGlu<td>99</td></td>	ValArgGlyPheSerAspLysAspLysLysLysTyrrPheTyrllysPhePheArgAspGlu <td>99</td>	99
Dd <td>1102<td>ATCTCTGGGTTCCTCGAGGCCCAAAGAAAGAGTACTTCTTCAAGTACTTCTCTGTGATGAG<td>116</td></td></td>	1102 <td>ATCTCTGGGTTCCTCGAGGCCCAAAGAAAGAGTACTTCTTCAAGTACTTCTCTGTGATGAG<td>116</td></td>	ATCTCTGGGTTCCTCGAGGCCCAAAGAAAGAGTACTTCTTCAAGTACTTCTCTGTGATGAG <td>116</td>	116
Qy <td>100<td>ArgArgAlaGluArgAlaTyrrArgPheValLysGluAsnGluThrLeuPheAlaLeuCys<td>119</td></td></td>	100 <td>ArgArgAlaGluArgAlaTyrrArgPheValLysGluAsnGluThrLeuPheAlaLeuCys<td>119</td></td>	ArgArgAlaGluArgAlaTyrrArgPheValLysGluAsnGluThrLeuPheAlaLeuCys <td>119</td>	119
Dd <td>1162<td>GCCCAAGCGCGGCGAGCCTTCAGTCTCATTCAGAGAAACAGAGGTCTCTTCCACCTGTGTC<td>122</td></td></td>	1162 <td>GCCCAAGCGCGGCGAGCCTTCAGTCTCATTCAGAGAAACAGAGGTCTCTTCCACCTGTGTC<td>122</td></td>	GCCCAAGCGCGGCGAGCCTTCAGTCTCATTCAGAGAAACAGAGGTCTCTTCCACCTGTGTC <td>122</td>	122
Qy <td>120<td>PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGly<td>139</td></td></td>	120 <td>PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGly<td>139</td></td>	PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGly <td>139</td>	139
Dd <td>1222<td>TTCATCCCCCTGGTCTGTGTGATCGTGTGCACTGGACTGAAACAGCAGATGGAGAGTGGC<td>128</td></td></td>	1222 <td>TTCATCCCCCTGGTCTGTGTGATCGTGTGCACTGGACTGAAACAGCAGATGGAGAGTGGC<td>128</td></td>	TTCATCCCCCTGGTCTGTGTGATCGTGTGCACTGGACTGAAACAGCAGATGGAGAGTGGC <td>128</td>	128
Qy <td>140<td>ArgAspLeuSerArgThrSerLysThrThrThrSerValTyrlleuLeuPheIleThrSer<td>159</td></td></td>	140 <td>ArgAspLeuSerArgThrSerLysThrThrThrSerValTyrlleuLeuPheIleThrSer<td>159</td></td>	ArgAspLeuSerArgThrSerLysThrThrThrSerValTyrlleuLeuPheIleThrSer <td>159</td>	159
Dd <td>1282<td>AAGAGCCTTGCCCGAGACATCTAAGACCAACCAACCGCGGTGTACGCTTCTTCTCCAGT<td>134</td></td></td>	1282 <td>AAGAGCCTTGCCCGAGACATCTAAGACCAACCAACCGCGGTGTACGCTTCTTCTCCAGT<td>134</td></td>	AAGAGCCTTGCCCGAGACATCTAAGACCAACCAACCGCGGTGTACGCTTCTTCTCCAGT <td>134</td>	134
Qy <td>160<td>ValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeu<td>179</td></td></td>	160 <td>ValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeu<td>179</td></td>	ValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeu <td>179</td>	179
Dd <td>1342<td>TGTGTCGACCCCGGGAGGAGCCAGGACACGGCTCTGTGCCCCCACTCTGGGGGGCTC<td>140</td></td></td>	1342 <td>TGTGTCGACCCCGGGAGGAGCCAGGACACGGCTCTGTGCCCCCACTCTGGGGGGCTC<td>140</td></td>	TGTGTCGACCCCGGGAGGAGCCAGGACACGGCTCTGTGCCCCCACTCTGGGGGGCTC <td>140</td>	140
Qy <td>180<td>CysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeu<td>199</td></td></td>	180 <td>CysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeu<td>199</td></td>	CysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeu <td>199</td>	199
Dd <td>1402<td>TCGCTCTTTGGTGGCAGATGTAATCTGGAACACGAAAAATCTGTTTGGAGAGTCCGACCTC<td>146</td></td></td>	1402 <td>TCGCTCTTTGGTGGCAGATGTAATCTGGAACACGAAAAATCTGTTTGGAGAGTCCGACCTC<td>146</td></td>	TCGCTCTTTGGTGGCAGATGTAATCTGGAACACGAAAAATCTGTTTGGAGAGTCCGACCTC <td>146</td>	146

```
QY 200 GluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeu 219
Db 2462 AGGAATCATGAGCTGCAGAAAGCGGATGTG---TCTGCTTTCCTGAGGATGAACCTGTTC 1518
QY 220 ProGlyValLeuGlnThrGluValThrGlnPheIleAspGlnSerPheGlnGluPhe 235
Db 1519 CAAAGGAAGTGGAGCTCGGAAGAGTTCTACAGCTTCATCCACATGACTTTCACAGGAGTTC 1578
QY 240 LeuAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGly 259
Db 1579 TTTGCGCCATGTACTACCTCTCGAAGAGAAAGGAAGGAGGACGAACGTTCCAGGG 1638
QY 260 -----ValGlyThrLeuLeuArgGlyAspAlaGln--- 269
Db 1639 AGTCCTTTGAAGCTTCCAGTCGAGAGCTGACAGTCTCTGGAAGAACTATGCAAAATTC 1698
QY 270 ProHisSerHisLeuValLeuThrArgPheLeuPheGlyLeuLeuSerAlaGluArg 289
Db 1699 GAAAGGGGTATTGATTTTGTGTGACGTTCTCTCTTGGCTTGSTAAACCGAGGAGG 1758
QY 290 MetArgAspIleGlnHisPheGlyCysValSerGluArgValLysGlnGlnAla 309
Db 1759 ACCTCCTACTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGCAATCAGGCTGGAGCTG 1818
QY 310 LeuArgTrpValGlnGlyGlnGlyGlnGlyCysProGlyValAlaProGluValThrGlu 329
Db 1819 CTGAATGGAT-----GAAAGTAAAGCC 1842
QY 330 GlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGluGlyGluGluPro 349
Db 1843 AAAGCTAAAGCTGCAG-----ATCAGCCAGCCAG----- 1875
QY 350 AsnTyrProLeuGluLeuLeuTyrCysLeuTyrGlnThrGlnGluAspAlaPheValArg 369
Db 1876 -----CTGGAATTTCTCTTTGATTGAGAACTGTCATCGGTGGAGTCACTGTCCCTGGG 1926
QY 370 GlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPheCysArgMetAsp 389
Db 1927 AGGGCCATGAGTATTTCCTCCCAAGATTGAGTCAAT-----CTCTCCACCAAGATGGAC 1980
QY 390 ValAlaValLeuSerTyrCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuIle 409
Db 1981 CACATGGTTTCTCTCTTTGATTGAGAACTGTCATCGGTGGAGTCACTGTCCCTGGG 2040
QY 410 SerCysArgLeuValAlaAlaGlnGlyLysLysLysSerLeuGlyLysArgLeuGln 429
Db 2041 TTCTCCATAACATGCCCAAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
QY 430 -----AlaSerLeuGlyGlyGlySerSerGlnGly---ThrTrp 441
Db 2101 ATGGTGCAGTGTGCTCCCAAGCTCTCTCATGCTGCCTGCTTCATGATTGGTGAAC 2160
QY 442 LysGlnLeuProAlaSerLeuHisProLeuPhe-----GlnAla 455
Db 2161 AGCCACCTCACTTCCAGTTTTGGCGGGGCTCTTTTCACTTGTAGCAGCAGCAGAGT 2220
QY 456 MetThr-----AspPro----- 459
Db 2221 CTAAGTGAATTGGACCTCAGTGACAATTTCTCTGGGGGACCCAGGAGTGAAGTGTGTGT 2280
QY 459 ----- 459
Db 2281 GAAAGCTCCAGCATCTCTGGCTGTAAACATTGGAGATTGTGGTGGGGCGCTGTGGCCTC 2340
QY 459 ----- 459
Db 2341 TCGCATGAGTGTGCTTGGATCTCTCTGCTGCTCCTCAGCAGCAACAGAGCTGGTGGAG 2400
QY 459 ----- 459
Db 2401 CTGGACCTGAGTGACAAGCCCTCGGTGACTTTCGGAATTCAGACTTCTGTGTGGGAGCTG 2460
```

```
QY 460 -----LeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAla 476
Db 2461 AAGCACCTGTGTGTCATCTGAAAGAGCTGTGTTGTGTCACTGCTGCTCCTCACATCAGCA 2520
QY 477 ValCysArgAspLeuSerGluAlaLeuArgAlaProAlaLeuThrGluLeuGlyLeu 496
Db 2521 TGTGTGAGGATCTTGATCAGTATTGAGCACCAGCATTCCCTGACCAGACTCTATGTG 2580
QY 497 LeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpPro 516
Db 2581 GGGGAGAAATGCTTGGGAGACTCAGGAGTGCATAATTTTATGTGAAAGAACCAAGATCCA 2640
QY 517 GlnCysArgValGlnThrValArgVal----- 525
Db 2641 CAGTGAACCTGCAGAAACTGGGTGTGTGAATTTCTGGCCTTACGTCTGTGTGTTC 2700
QY 525 ----- 525
Db 2701 GCTTTGCTCCTCGGTACTCAGCACAATCAGAAATCTCAGGCACCTTTACCTCGGAGGCAAC 2760
QY 526 GlnLeuProAspProGlnArgGlyLeuGlnTyrLeuVal----GlyMetLeuArgGlnSer 544
Db 2761 ACTCTCGGAGAC-----AAGGGGATCAAACTACTCTGTGTGAGGAGCTTTCACCCCGAC 2814
QY 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMet----- 561
Db 2815 TCGAAGCTTCAAGTGTGTGAATTAGACAACTGCAACCTCACACTGCTGTGTGTGGAT 2874
QY 561 ----- 561
Db 2875 CTTTCCACACTTCTGACCTCCAGCCAGAGCCTGGGAAAGCTGAGCCTGGGCAACAATGAC 2934
QY 562 -----ValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeu 576
Db 2935 CTGGCGAGCTGGGGGTTCATGATGTTCTGTAAGTGTGAAAGCTGAAACAGCAGAGCTGCCCTCTG 2994
QY 577 GlnThrLeuSerLeuAlaSerValGluLeuSerGlnGlnSerLeuGlnGluLeuGlnAla 596
Db 2995 CAGAACCTGGGGTGTCTGAAATGATTTCAATTATGAGACAAAAAGTGCCTTAGAAACA 3054
QY 597 ValLysArgAlaLysProAspLeuValIle 606
Db 3055 CTTCAAGAAGAAAGCCTGAGCTGACCGTC 3084
RESULT 5
US-10-127-516-6
; Sequence 6, Application US/10127516
; Publication No. US20020187922A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-370001
; CURRENT APPLICATION NUMBER: US/10/127,516
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 10/027,629
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 09/964,955
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506, 067
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-127-516-6
```

Alignment Scores:

Pred. No.: 3,35e-81 Length: 3102
 Score: 800.50 Matches: 219
 Percent Similarity: 43.64% Conservatives: 117
 Best Local Similarity: 28.44% Mismatches: 241
 Query Match: 24.88% Indels: 193
 DB: 13 Gaps: 17

US-10-029-347-2 (i-625) x US-10-127-516-6 (1-3102)

QY	1	MetLeuAlaGlnProGlnArgLeuLeuPheLeuLeuAspGlyAlaAspGluLeuProAla	20
DB	862	ATCGTGAGAAAAACCCCTCCAGATCTCTTCTCATGAGCGGTTCGATGAGTGAAGT	921
QY	21	LeuGlyGlyProGluAlaAlaPro---CysThrAspProPheGluAlaAlaSerGlyAla	39
DB	922	GCCTTTGACGACACATAGACCGCTCTCCATGACTGGCAGAGCGCCAGCGGGAGAC	981
QY	40	ArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr	59
DB	982	ATTCTCTGAGCAGCCTCATCAGAAAGAGCTGCTTCCGAGGCTCTCTGCTCATCACC	1041
QY	60	ThrArgAlaAlaAlaProGlnArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu	79
DB	1042	ACGAGACCTGTGCGCCCTCGAGAAACTGACGACTTCTGAGACCATCTCGCATGTGGAG	1101
QY	80	ValArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLysLys	99
DB	1102	ATCCTGGTCTTCCGAGGCCAAAGGAAGAGTACTTCTCAAGTACTTCTCTGATGAG	1161
QY	100	ArgArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCys	119
DB	1162	GCCAAGCCAGGCGCAGCCTTCAGTCTGATGATGAGGAGGAGGAGGAGGAGGAGGAG	1221
QY	120	PheValProPheValCysTrpIleValCysThrValLeuArgGlnGluLeuGluLeuGly	139
DB	1222	TTATCTCCCTCGTCTGCTGGATCGGTGCTGATGAGTGAAGTGAAGTGAAGTGAAGT	1281
QY	140	ArgAspLeuSerArgThrSerLysThrThrThrThrThrThrThrThrThrThrThrThr	159
DB	1282	NAGAGCCTTGGCCAGACATCTAAGACCACCCAGCGGTGATGCTTCTTCTTTCAGT	1341
QY	160	ValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeu	179
DB	1342	TTGCTGACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1401
QY	180	CysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeu	199
DB	1402	TGCTCTTTGGCTGCAGATGGAATCTGGAAACCCAGAAATCTCTTTGAGGAGTCCGACCTC	1461
QY	200	GluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeu	219
DB	1462	AGGAATCATGGACTGCAGAAGCGGATGTG---TCTGCTTCTCTGAGGATGAACCTGTTC	1519
QY	220	ProGlyValLeuGluThrCysValThrThrThrThrThrThrThrThrThrThrThrThr	239
DB	1519	CAAAAGGAAGTGGAGCTGGCAGAGAGTCTACAGCTTCATCCACATGACTTCCAGGAGTTC	1578
QY	240	LeuAlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGly	259
DB	1579	TTTGCCGCCATGTACTACTCTGTGGAAGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1638
QY	260	-----ValGlyThrLeuLeuArgGlyAspAlaGln---	269
DB	1639	AGTCGTTTGAAGCTTCCAGCGGAGCGTGACAGCTCTTCTGGAAACATATGGCAATTC	1698
QY	270	ProHisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArg	289
DB	1699	GAAAAGGGGTATTTGATTTTGTGTACGTTTCTCTTTGGCCCTGGTAAACAGGAGGAGG	1758
QY	290	MetArgAspIleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAla	309
DB	1759	ACCTCTACTTGGAGAGGAAGAAATTAAGTTGCCAAGATCTCTCAGCAAAATCAGGCTGGAGCTG	1818

QY	310	LeuArgTrpValGlnGlyGlnGlnGlyCysProGlyValAlaProGluValThrGlu	329
DB	1919	CTGAATGGAATT-----CAAGTGAAGGCC	1842
QY	330	GlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGluGluGluGluGluPro	349
DB	1843	AAAGCTAAAGAGCTGCAG-----ATCCAGCCAGCCAG-----	1875
QY	350	AsnTyrProLeuLeuLeuLeuTyrCysLeuTyrGluThrGlnGluAspAlaPheValArg	369
DB	1876	-----CTGAAATTTGTTACTCTTTTGTACGATCGAGGAGGAGGACTTCGTGCAA	1926
QY	370	GlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPheCysArgMetAsp	389
DB	1927	AGGCCATGACACTATTTCCCAAGATTGATGATCAAT-----CTCTCCACCAATGGAC	1980
QY	390	ValAlaValLeuSerTyrCysValArgCysProAlaGlyGlnAlaLeuArgLeuIle	409
DB	1981	CACATGTTTCTTCTTCTTTCATTTGAGACTGTCTCATCGGTGAGTCACTGTCTCTGGG	2040
QY	410	SerCysArgLeuValAlaAlaGlnGluLysLysLysLysLysLysLysLysLysLysLys	429
DB	2041	TTTCTCATAACATGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2100
QY	430	-----AlaSerLeuGlyGlyGlySerSerGlnGly---ThrThr	441
DB	2101	ATGCTGAGTGTCTCTCCCAAGCTCTCTCATCTGCTGCTGTTCTCATGGATTTGTTGAAC	2160
QY	442	LysGlnLeuProAlaSerLeuLeuHisProLeuPhe-----GlnAla	455
DB	2161	AGCCAGCTCATCTCCAGTCTTTTCCGCGGCGCTCTTTTCAGTTCTGAGCACCAGCAGAGT	2220
QY	456	MetThr-----AspPro-----	459
DB	2221	CTAACTGAAATTGGACCTCAGTGACAAATCTCTGGGGGACCCAGGGATGATGAGTGTGTGT	2280
QY	459	-----	459
DB	2281	GAAACGCTCCAGCATCTCGCTGTACATTCGAGAGATTGTGTTGGGCGCTGTGSCCTC	2340
QY	459	-----	459
DB	2341	TCGCATGAGTGTGCTTCGACATCTCTTGGTCTCAGCAGCAACAGAAAGTGTGTGAG	2400
QY	459	-----	459
DB	2401	CTGGACCTGAGTGACAAACCCCTCGGTGACTTCGGAAATCAGACTTCTGTGTGTGGAGCTG	2460
QY	460	-----LeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAla	476
DB	2461	AASACCTGCTGTGTGCAATCTGAAGAAGCTCTGTTGGTCTGCTGCTCATCATCAGCA	2520
QY	477	ValCysArgAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeu	496
DB	2521	TGTTGTGAGATCTTCATCAGTATTGAGCACAGCCATTCTCTGACGACTCTATGTS	2580
QY	497	LeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGlyLeuAlaLysPro	516
DB	2581	GGGAGAAATGCTTTGGGAGCTCAGGAGTGGCAATTTTATGTGAAAACCCCAAGAAATCCA	2640
QY	517	GlnCysArgValGlnThrValArgVal-----	525
DB	2641	CAGTGTAACTGCAGAAACTGGGGTGGTGAATTCCTGGCTTACCTCAGTCTGTTGTTCA	2700
QY	525	-----	525
DB	2701	GCTTTGCTCGGTACTCAGCACTAATCAGAAATCTCAGCACTTTTACCTGCGAGGCAAC	2760
QY	526	GlnLeuProAspProGlnArgGlyLeuGlnTyrLeuVal---GlyMetLeuArgGlnSer	544
DB	2761	ACTCTCGGAGAC-----AAGGGATCAAACTACTCTGTGAGGAGTCTTTCACCCCGAC	2814
QY	545	ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMet-----	561


```
Qy 442 LysGlnLeuProHisSerLeuHisProLeuPhe-----GlnAla 455
Db 2161 AGCCACCTCAGTCCAGTATTTGCCGGGCGCTTTTCAGTCTCGAGCACCAGCCAGAGT 2223
Qy 456 MetThr-----AspPro----- 459
Db 2221 CTAACTGAATGAGCTCAGTGACATTTCTGGGGGACCCAGGATGAGAGTGTGTGT 2280
Qy 459 ----- 459
Db 2281 GAAACGCTCCAGCATCTGGCTGTAACTTCGGAGATTGTGTGGGCGCTGTGGCCCTC 2340
Qy 459 ----- 459
Db 2341 TCGCATGAGTGTCTTGGACATCTCTCTGGTCTCAGCAGAACAGAGCTGTGGAG 2400
Qy 459 ----- 459
Db 2401 CTGGACCTGAGTGACACGCCCTCGCTGACTTCGGAAATCAGACTTCTGTGTGGGACTG 2460
Qy 460 -----LeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAla 476
Db 2461 AAGCAGCTGTGTGCAATCTGAAGAAGCTCTGTGTGCTGAGTGTGTCTCATCATCAGCA 2520
Qy 477 ValCysArgAspLeuSerGlnAlaLeuArgAlaProAlaLeuThrGlnLeuGlyLeu 496
Db 2521 TGTGTGAGGATCTTGATCAGTATTGAGCAGCAGCATTTCTGACACAGACTCTATGTG 2580
Qy 497 LeuHisAspArgLeuSerGlnAlaGlyLeuArgMetLeuSerGlnGlyLeuAlaTrpPro 516
Db 2581 GGGAGAAATGCCCTGGGAGACTCAGGAGTCGCAATTTATGTGAAAAGCCCAAGATCCA 2640
Qy 517 GlnCysArgValGlnThrValArgVal----- 525
Db 2641 CAGTGTAACTGCAGAAATCGGGTGTGGTAAATCTGGCCTTACGTCAGTCTGTGTCTTCA 2700
Qy 525 ----- 525
Db 2701 GCCTTGTCTCGTACTCAGCACTAATCAGAATCTCAGGCACCTTTACCTGCGAGSCAAC 2760
Qy 526 GlnLeuProAspProGlnArgGlyLeuGlnTyrLeuVal---GlyMetLeuArgGlnSer 544
Db 2761 ACTCTCGGAGC-----AAGGGATCAAACTACTCTGTGAGGAGACTTTCACCCCGAC 2814
Qy 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMet----- 561
Db 2815 TGCAGCTTCAGGTGTGGAAATTAGACAACTGCAACTCAGTCACACTGCTGCTGGGAT 2874
Qy 561 ----- 561
Db 2875 CTTTCCACACTTCTGACCTCCAGCAGAGCCTCGAAAGCTGAGCCTGGGCAACATGAC 2934
Qy 562 -----ValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeu 576
Db 2935 CTGGGCACTGGGGGTATGATGTCTGTGAAGTGTCTGAAACAGCAGAGCTGCTCTCTG 2994
Qy 577 GlnThrLeuSerLeuAlaSerValGlnLeuSerGlnGlnSerLeuGlnLeuGlnAla 596
Db 2995 CAGAACTGGGGTGTCTGAAATGATTTTCAATATGAGACAAAAGTCGGTTGAGAAACA 3054
Qy 597 ValArgAlaLysProAspLeuValle 606
Db 3055 CTTCAAGAAAAGCCTGAGCTGACCGTC 3084

RESULT 7
US-10-132-967-4
; Sequence 4, Application US/10132967
; Publication No. US20030170841A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-311001
```

```
; CURRENT APPLICATION NUMBER: US/10/132,967
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US/09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506,067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3857
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-967-4

Alignment Scores:
Pred. No.: 4,66e-81 Length: 3857
Score: 800.50 Matches: 219
Percent Similarity: 43.64% Conservative: 117
Best Local Similarity: 28.44% Mismatches: 241
Query Match: 24.88% Indels: 193
DB: 12 Gaps: 17

US-10-029-347-2 (1-625) x US-10-132-967-4 (1-3857)
Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
Db 1000 ATCGTGAGAAACCTCCAGAAATCTCTCTCATGACGGCTTCGATGAGCTGCAAGGT 1059
Qy 21 LeuGlyGlyProGluAlaAlaPro---CysThrAspProPheGluAlaAlaSerGlyAla 39
Db 1060 GCCTTTGACGAGCACATAGGACGCTCTGCATCTGACTGACAGAGGCGCGGGGAGAC 1119
Qy 40 ArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 59
Db 1120 ATCTCTCAGCAGCCTCATCAGAAAGAGCTCTCTCCGAGCCCTCTCTGCTCATCACC 1179
Qy 60 ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu 79
Db 1180 ACGAGACCTGTGGCCCTGGAGAAACTGCAGCACTTGTGTGACCATCTCGGCATGTGGAG 1239
Qy 80 ValArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLys 99
Db 1240 ATCTCTGGTCTTCCGAGGCCAAAGAAAGAGTACTTCTTCAAGTACTTCTCTGATGAG 1293
Qy 100 ArgArgAlaGluArgAlaTyrArgPheValLysGluAsnGlnThrLeuPheAlaLeu 119
Db 1306 GCCCAACCCAGGACGCTTCTGCTGATTGAGAGACGAGTCTCTTCCACCATGTGC 1359
Qy 120 PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGly 139
Db 1360 TTCAATCCCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1419
Qy 140 ArgAspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSer 159
Db 1420 AAGAGCCTTCCAGACATCTAAGACCCACCCGCGGTGTACGTCTTCTTCTTCTTCTCAGT 1479
Qy 160 ValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeu 179
Db 1480 TTGCTGACGCCCGGGGAGGAGCCAGGACGCGCTCTGCGCCACCTCTCGGGGCTC 1539
Qy 180 CysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeu 199
Db 1540 TGCTCTTTGCTGACATGGAATCTGAAACCAACCAAAATCTGTTTGGAGAGTCGACCTC 1599
Qy 200 GlnGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeu 219
Db 1600 AGGAATCATGAGTGCAGAAAGCGGATGTG---TCTGCTTCTCTGAGGATGAACCTGTTC 1656
Qy 220 ProGlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPhe 239
Db 1657 CAAAGGAAGTGGAGTCCGAGAGATTTCTACAGCTTCTATCCACATGACTTCTCCAGGAGTTC 1716
Qy 240 LeuAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGly 259
```

```
Db 1717 TTTGCCCCATGACTACTACCTGCTGGAAGAAAGAGAGCAACGTTCCAGGG 1776
QY 260 -----ValGlyThrLeuLeuArgGlyAspAlaGln--- 269
Db 1777 AGTCGTTTGAAGCTTCCAGCGAGACGTCACATCTCTGGAAACTACTGCAAAATTC 1836
QY 270 ProHisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAAGluArg 289
Db 1837 GAAAAGGGTATTCATTTTGTGTACGTTTCTCTTGGCCTGGTAAACCCAGGAGGG 1896
QY 290 MetArgAspIleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAla 309
Db 1897 ACCTCCTACTTGGAGAGAAATTAAGTTGTCAGATCTCTCAGCAANTCAGCGCTGGAGCTG 1956
QY 310 LeuArgTrpValGlnGlyGlnGlnGlyCysProGlyValAlaProGluValThrGlu 329
Db 1957 CTGAATGGATT-----GAAGTAAAGCC 1980
QY 330 GlyAlaCysGlyLeuGluGluProGluGluGluGluGluGluGluGluPro 349
Db 1981 AAAGCTAAAGCTGCAG-----ATCCAGCCCGCCAG----- 2013
QY 350 AsnTyrProLeuGluLeuLeuTyrCysLeuTyrGluThrGlnGluAspAlaPheValArg 369
Db 2014 -----CTGGAATTCCTTACTGTTTGTACGATGCGAGGAGGAGGACTTCGCGAA 2064
QY 370 GlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPheCysArgMetAsp 389
Db 2065 AGGCCATGGACTATTTCCTCCCAAGATGAGTCAAT-----CTCTCCACCAAGATGGAC 2118
QY 390 ValAlaValLeuSerTyrCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuLe 409
Db 2119 CACATGGTTTCTTCTTGTGATTGAGAACTGTCATCGGTGGAGTCACTGTCCTCGGG 2178
QY 410 SerCysArgLeuValAlaAlaGlnGluLysLysLysSerLeuGlyLysArgLeuGln 429
Db 2179 TTTCTCCATACATGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2238
QY 430 -----AlaSerLeuGlyGlyGlySerSerGlnGly---ThrThr 441
Db 2239 ATGTTGCAGTGTGCTCCCAAGCTCTCTCATGCTGCTGCTTCTCATGGATTGGTGAAC 2298
QY 442 LysGlnLeuProAlaSerLeuLeuHisProLeuPhe-----GlnAla 455
Db 2299 AGCCACCTCACTTCAGTTTTTGGCGGCGCTCTTTTCAGTTCTGAGCACCAGCCAGAT 2358
QY 456 MetThr-----AspPro----- 459
Db 2359 CTAAGTGAATTGGACCTCAGTGACAATTTCTCTGGGGACCCAGGGATGAGAGTCTGTGT 2418
QY 459 ----- 459
Db 2419 GAAACGCTCCAGCATCTCTGGCTGTAAACATTCGGAGATTGTGTTGGGGCGCTGTGGCTC 2478
QY 459 ----- 459
Db 2479 TCGCATGAGTCTGCTTCGACATCTCTTGTGCTCTCAGCACAACCAAGAGCTGTGGAG 2538
QY 459 ----- 459
Db 2539 CTGGACCTGAGTGACAACGCCCTCGGTGACTTCGGAAATCAGACTCTCTGTGTGGGACTG 2598
QY 460 -----LeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAla 476
Db 2599 AAGCACCTGTTGTGCAATCTGAAGAGCTCTGGTTGGTCAGCTGCTGCTCACAATCAGCA 2659
QY 477 ValCysArgAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeu 496
Db 2659 TGTGTGAGGATCTTGATCAGTATGAGCACCAGCCATTCCTGACCCAGACTTATGTGT 2719
QY 497 LeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaLeuPro 516
```

```
Db 2719 GGGGAGAATGCTTGGGAGACTCAGAGTCGCAATTTTATGCGAAAAGCCAGAAATCCA 2778
QY 517 GlnCysArgValGlnThrValArgVal----- 525
Db 2779 CAGTGTAACTGCAGAAACTGGGGTTGGTGAATCTGGCCTTACGTCAGTCTGTGTTTCA 2838
QY 525 ----- 525
Db 2839 GCTTTGCTCGTGTACTCAGCACTAATCAGAATCTCAGCACCTTTTACCTGGGAGGCAAC 2899
QY 526 GlnLeuProAspProGlnArgGlyLeuGlnTyrLeuVal---GlyMetLeuArgGlnSer 544
Db 2899 ACTCTCGGAGAC-----AAGGGATCAAACTACTCTGTGAGGAGACTCTTGCAACCCGAC 2952
QY 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMet----- 561
Db 2953 TGCAGACTTCAGTGTGTGGAAATTAGACAACCTCAACCTCAGCTCACACTGCTGTGGAT 3012
QY 561 ----- 561
Db 3013 CTTTCCACACTTCTGACCTCCAGCCAGAGCCTGGCAAGCTGAGCCTGGGCAACAAATGAC 3072
QY 562 -----ValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeu 576
Db 3073 CTGGCGACCTGGGGTTCATGATGTTCTGTGAAGTGTGAAACAGCAGAGCTGCTCTGTG 3132
QY 577 GlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnLeuGlnAla 596
Db 3133 CAGAACCTGGGGTGTCTGGAATGATTTCAATTATGAGACAAAGAGTGGTTAGAAACA 3192
QY 597 ValLysArgAlaLysProAspLeuValIle 606
Db 3193 CTTCAAGAGAAAGCCTGAGCTGACCGCTC 3222
```

RESULT 9

```
US-10-127-516-4
; Sequence 4, Application US/10127516
; Publication No. US20020187922A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Manji, Gulam A.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN
; FILE REFERENCE: 07334-370001
; CURRENT APPLICATION NUMBER: US/10/127,516
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 10/027,629
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 09/964,955
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506, 067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (139)...(3240)
US-10-127-516-4
```

```
Alignment Scores:
Pred. No.: 4,66e-81 Length: 3857
Score: 800.50 Matches: 219
Percent Similarity: 43.64% Conservative: 117
Best Local Similarity: 28.44% Mismatches: 241
Query Match: 24.88% Indels: 193
DB: 13 Gaps: 17
```



```
QY 562 -----ValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeu 576
Db 3073 CTGGGCGACCTGGGGGTCATGATCTTGTGAAGTGTCTGAACAGCAGACGTCCTCTG 3132
QY 577 GlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnGluLeuAla 596
Db 3133 CAGAACCTGGGGTGTCTGANATGTAATTCATATATGAGACAAAAGTGGGTAGAACAA 3132
QY 597 VallysArgAlaLysProAspLeuValle 606
Db 3193 CTTCAAGAGAAAGCCTGA3CTGACCGTC 3222
RESULT 9
US-10-027-629-4
; Sequence 4, Application US/10027:29
; Publication No. US20020197660A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Masji, Gulam A.
; TITLE OF INVENTION: NOVEL MO-ECULES OF THE PYRIN DOMAIN
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-341001
; CURRENT APPLICATION NUMBER: US/10/027,629
; CURRENT FILING DATE: 2001-12-2
; PRIOR APPLICATION NUMBER: US 09/964,955
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506,067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEC ID NO 4
; LENGTH: 3857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (139)...(3240)
US-10-027-629-4
Alignment Scores:
Pred. No.: 4.66e-81 Length: 3857
Score: 800.50 Matches: 219
Percent Similarity: 43.64% Conservative: 117
Best Local Similarity: 28.44% Mismatches: 241
Query Match: 24.88% Indels: 193
Gaps: 17
US-10-029-347-2 (1-625) x US-10-027-629-4 (1-3857)
QY 1 MetLeuAlaGlnProGlnArgLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
Db 1000 ATCGTGAAACCCCTCCAGAACTCTCTCTCATGACGCGCTTCGATGAGCTGCAAGGT 1059
QY 21 LeuGlyGlyProGluAlaAlaPro---CysThrAspProPheGluAlaAlaSerGlyAla 39
Db 1060 GCCTTTGACGAGCACAATAGGACCGCTCTGCACTGCTGCGAGAGGCGGCGGAGAC 1119
QY 40 ArgValLeuGlyGlyLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 59
Db 1120 ATTCTCTGAGCAGCCTCATCAGAAAGAGCTGCTTCCGAGGCGCTCTGCTCATCACC 1179
QY 60 ThrArgAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu 75
Db 1180 ACGAGACTGTGGCCCTGGAGAACTGCGACACTGTGTCGACCATCTCTGCGATGTGGAG 1239
QY 80 ValArgGlyPheSerAspLysAspLysLys---ysTyrPheTyrLysPhePheArgAspGlu 99
Db 1240 ATCTGGGGTTCTCGAGGCCAAAAGAAAGAGACTTCTTCAAGTACTTCTCTGATGAG 1299
QY 100 ArgArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCys 119
Db 1000 ATCTGGGGTTCTCGAGGCCAAAAGAAAGAGACTTCTTCAAGTACTTCTCTGATGAG 1299
```

```
Db 1300 GCCCAAGCCAGGGGAGCCTTCAGTCTGATTGAGAGAACAGGTCCTCTTCCACCATGTGC 1359
QY 120 PheValProPheValCysTyrIleValCysThrValLeuArgGlnGlnLeuGluLeuGly 139
Db 3360 TTCATCCCTGGTCTGCTGATCGTGTGCACACTGCACTGAAACAGCAGCATGAGAGTGGC 3419
QY 140 ArgAspLeuSerArgThrSerLysThrThrSerValTyrLeuLeuPheIleThrSer 159
Db 3420 AAGAGCCTTGCCAGACATCTAAGACCAACCCCGGGGTGATCGTCTTCTCTTTCAGT 1479
QY 160 ValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeu 179
Db 1480 TTGCTGCAGCCCGGGGAGGAGCAGGAGCAGCGCTCTCGCCACCTCTGGGGGCTC 1539
QY 180 CysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGlyLysGluLeu 199
Db 1540 TGCTCTTTGGCTGCAGATGGAATCTGGAAACAGAAATCTCTGTTGAGAGAGTCCGACCTC 1599
QY 200 GluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeu 219
Db 1600 AGCAATCATGACTGCAGAGGCGGATGTG---TCTGCTTCTTGAGGATGAACCTGTTC 1656
QY 220 ProGlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPhe 239
Db 1657 CAAAAGGAAGTGGACTGCGAGAAGTTCTACAGCTTCATCCACATGACTTCCAGGATTC 1716
QY 240 LeuAlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGly 259
Db 1717 TTTGCGCCATGTACTACTCTGGAAGAGAAAGGAAGGAGGAGCAAGCTTCCAGG 1776
QY 260 -----ValGlyThrLeuLeuArgGlyAspAlaGln--- 269
Db 3777 AGTCGTTTGAAGCTTCCCGAGCGAGCGTGCAGCTCTCTGGAAGAACTATGGCAAAATC 1836
QY 270 ProHisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArg 289
Db 1837 GAAAAGGGGTATTTGATTTTGTGTTTGTCTCTTTGGCTGTGTAACAGCAGGAGG 1896
QY 290 MetArgAspIleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAla 309
Db 1897 ACCTCCTACTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGCAATCAGGCTGGAGCTG 1956
QY 310 LeuArgTyrValGlnGlyGlnGlyCysProGlyValAlaAlaProGlyValThrGlu 329
Db 1957 CTGAAATGGAAT-----GAAAGTGAAGGCC 1983
QY 330 GlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGluGlyGluGluPro 349
Db 1981 AAAAGCTAAAAGCTGCAG-----ATCCAGCCCGCCAG 2013
QY 350 AsnTyrProLeuLeuLeuLeuTyrCysLeuTyrGluThrGlnGluAspAlaPheValArg 369
Db 2014 -----CTGGAATTTCTTCTGTTTGTACGAGATGCGAGGAGGAGGACTTCGTGCA 2064
QY 370 GlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPheCysArgMetAsp 389
Db 2065 AGGGCCATGAGCTATTTCCCAAGATTGAGATCAAT-----CTCCACCAAGATGCA 2118
QY 390 ValAlaValLeuSerTyrCysValArgCysProAlaGlyGlnAlaLeuArgLeuIle 409
Db 2119 CACATGGTTTCTCTCTTTTGGCATGAGAACTGTATCGGGTGGAGTCACTGCTCGTGGG 2179
QY 410 SerCysArgLeuValAlaAlaGlnGluLysLysLysSerLeuGlyLysArgLeuGln 429
Db 2179 TTTCTCCATACATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2238
QY 430 -----AlaSerLeuGlyGlySerGlnGly---ThrThr 441
Db 2239 ATGGTGCAGTGTCTCTCCCAAGCTCTCTCATGTGCTGTCTCTCATGGATTGGTGAAC 2298
QY 442 LysGlnLeuProAlaSerLeuLeuHisProLeuPhe-----GlnAla 455
Db 2299 AGCCACCTCACTTCCAGTTTTCGCGGGGCTCTTTTCAGTTCTGAGCACCAGCAGAGT 2358
```

```
QY 456 MetThr-----AspPro----- 459
D 2359 CTAACCTGAATTGGACCTCAGTCACAACTCTCTGGGGACCCAGGATGAGAGTGTGTG 2418
QY 459 ----- 459
D 2419 GAAACGCTCCAGCATCTCGCTGTAACATTCGAGATTGTGTGGGGGCTGTGGCCTC 2478
QY 459 ----- 459
D 2479 TCGCATGAGTGTCTTCGACATCTCTTGGTCTCAGCAGCAACAGAGCTGTGGAG 2538
QY 459 ----- 459
D 2539 CTGGACCTGAGTGACAAACCCCTCGGTGACTTCGGAATCAGACTTCTGTGTGGGACTG 2598
QY 460 -----LeuHisLeuSerSerLeuThrLeuSerHisCysLeuProAspAla 476
D 2599 AAGCACCTTGTGTGCAATCTGAAGAAGCTCTGTGTGTGCTGCTCAGATCAGCA 2658
QY 477 ValCysArgAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeu 496
D 2659 TTTGTGTCAGGATCTTCATCAGTATTCAGCAGCCATCTCCTGACAGACTCTATGTG 2718
QY 497 LeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpPro 516
D 2719 GGGGAGATGCTCTGGGAGACTCAGAGTCGCAATTTATGTGAAAGCCAGATCCA 2778
QY 527 GlnCysArgValGlnThrValArgVal----- 525
D 2779 CAGTGTAACTGCAGAAACTGCGGTGTGTGAATTTCTGGCCTTACGTCAGTCTGTGTTCA 2838
QY 525 ----- 525
D 2839 GCTTTGTCTCTCGGTACTCAGCAGCTAATCAGAACTCAGCACCTTTACCTCGAGGCAAC 2898
QY 526 GlnLeuProAspProGlnArgGlyLeuGlnTyrLeuVal---GlyMetLeuArgGlnSer 544
D 2899 ACTCTCGGAGAC-----AAGSGGATCAAACTACTCTGTGAGGACTCTTGACCCCGAC 2952
QY 545 ProAlaLeuThrLeuAspLeuSerGlyCysGlnLeuProAlaProMet----- 561
D 2953 TCGAAGCTTCAGGTGTGTGAATTAGAACAACTGCAACCTCAGCTCACACTGTCTGGGAT 3012
QY 561 ----- 561
D 3023 CTTTCCACACTTCTGACCTCCAGCCAGAGCTCGGAAAGCTGAGCTGGGCAACAATGAC 3072
QY 562 -----ValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeu 576
D 3073 CTGGGCGACCTGGGGTCTATGATGTTCTGTGAAGTGTGAAACAGCAGAGCTGCTCTCG 3132
QY 577 GlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnGlnAla 596
D 3133 CAGAACCTCGGGTGTCTGAAATGATTTTCAATTTATGAGACAAAGTGCCTTAGAACA 3192
QY 597 ValLysArgAlaLysProAspLeuValille 606
D 3193 CTTCAAGAAAGAAAGCCCTGAGTGACCGTTC 3222
RESULT 10
US-09-996-617-5
; Sequence 5, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/03/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
```

```
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-996-617-5

Alignment Scores:
Pred. No.: 1,45e-76 Length: 4287
Score: 762.00 Matches: 220
Percent Similarity: 48.79% Conservative: 103
Best Local Similarity: 33.23% Mismatches: 207
Query Match: 23.68% Indels: 132
DB: 10 Gaps: 22

US-10-029-347-2 (1-625) x US-09-996-617-5 (1-4287)
QY 1 MetLeuAlaGlnProGlnArgLeuPheLeuPheLeuAspGlyLeuPro--- 19
D 1192 ATCTGTCTAGGCACAGCGGCTGCTTTCATCTCGATGGTGTAGATGAG---CCAGGA 1248
QY 20 ---AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaSerGly 38
D 1249 TGGGTCTTGCAGGCGGAGTCTGTGCTCTGCTGAGCCGACGACGCGCGG 1308
QY 39 AlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuProThrAlaLeuLeuVal 58
D 1309 GATGCACTGTCTGGGAGTTTGTCTGGGAAAACATATCTCCGAGGATCTTCTTGTATC 1368
QY 59 ThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla 78
D 1369 ACGCTCGGACACAGCTCTGCAGAACCTTCTCTTCTTGGAGCAGGACGCTTGGTA 1428
QY 79 GluValArgGlyPheSerAspLysLysLysLysLysLysLysLysLysLysLysLysLys 98
D 1429 GAGGTCTCTGGGTTCTCTGAGTCCAGCAGGAAAGAAATATTTCTACAGATATTTTACAGAT 1488
QY 99 GluArgArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeu 118
D 1489 GAAAGGCAACAATTAGAGCCTTTAGTGGTGGTCAATCAAAACAAAGAGCTCTGGGCGCTG 1548
QY 119 CysPheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeu 138
D 1549 TGTCTTGTGGCTGGGTGTCTCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1608
QY 139 GlyArgAspLeuSerArgThrSerLysThrThrThrThrThrThrThrThrThrThrThr 158
D 1609 AAGGAAATCTACACTGACTTCCAGACCCACACCAACCTCTGTCTACATTTTACAGAT 1668
QY 159 SerValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsn 178
D 1669 CAGCTCTCCAGCTCAGCAATG-----GGACCCGAG-----CTCAGAGAC 1710
QY 179 LeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGlu 198
D 1711 CTCTGTCTCTGGCTGTCTGAGGCGCATCTGCCAAAAAAGACCCCTTTTTCAGTCCAGATGAC 1770
QY 199 LeuGluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGlu 218
D 1771 CTCAGGAAGCATGGGTTAGATGGGGCCATCTCTCCACC---TTCTTTAAGATG----- 1821
QY 219 LeuProGlyValLeuGluThrGlu-----ValThrTyrGlnPheIleAspGlnSer 235
D 1822 -----GGTATTTCTTCAAGAGACCCCATCCCTCTGTGAGCTAGAGCTTCAATTCACCTCTGT 1875
QY 236 PheGlnGluPheLeuAlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThr 255
D 255 ----- 255
```

```
Db 1876 TTCCAAAGAGTCTTTGGCAGCATCTCTATGCTTGGAGGAT-----:917
Cy 256 AialaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuVal 275
Db 1918 --CAGAGGGGAGAGGT-----AAACATTCTAAATTGCATC 1950
Cy 276 Leu-----Thr 277
Db 1951 ATAGATTGGAAAGACGCTAGAGCATATGGATATACATGGCCCTGTTGGGGCATCAACC 2010
Cy 278 ThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPhe 297
Db 2011 ACACGTTTCTCTATTGGGCTCTTAAGTCATGAGGGGAGAGAGATGGAGACATCTTT 2270
Cy 298 GlyCysMetValSerGluArgValGlyGlnAlaLeuArgTrpValGlnGlyGlnGly 317
Db 2071 CACTGCCGGCTGCTCAG-----GGGAGGAACTGATGCGATGGGTCT----- 2112
Cy 318 GlnCysProGlyValAlaProGluValThrGluGlyAlaCysGlyLeuGlyAspThr 337
Db 2113 -----CGTCCCTGCAGCTGCTG 2130
Cy 338 GluGluProGluGluGluGluGluGluGluProAsnTyrProLeuGluLeuLeuTyr 357
Db 2231 CTGCAGCCA-----CACTCTCTGGAGTCCCTCCAC 2160
Cy 358 CysLeuTyrGluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGlu 377
Db 2262 TGCTTTACGAGACTCGGAACAAACCTTCTGTACACAAAGTATGGCCCATTTGGAAGAA 2220
Cy 378 LeuAlaLeuGlnArgValArgPheCys-----ArgMetAspValAlaValLeuSer 394
Db 2222 ATGGGCATG-----GTGTAGAACACACATGGAGCTCTTAGTGTGCAC 2265
Cy 395 TyrCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuLeuSerCysArg----- 412
Db 2266 TCTGCAATTAATTCAGCCGTCACGTGAAGAAGCTTCAGCTGATTGAGGGCAGGCAGCAC 2325
Cy 412 ----- 412
Db 2326 AGATCAACATGGAGCCGCCACATGGTGTAGTCTGTTACAGTGGGTCCACGATCAGATGCC 2395
Cy 413 -----LeuValAlaAlaGlnGluGlyGlySerLeuGlyGlyArgLeuGln 429
Db 2386 TATTGGCAGATTCTTCTCTCCTCAGCTCACAGAACCTG---AAGGAGCTGGAC 2442
Cy 430 AlaSerLeuGlyGlyGlySerSerGlnGlyThrThrLysGlnLeuProAlaSerLeuLeu 449
Db 2443 CTAAGT-----GGAACCTCG-----CTGAGCCACTCTTCGAGTG 2475
Cy 450 HisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSerSerLeuThrLeuSer 469
Db 2476 AAGACTCTTTGTAGACCTTGAAGCCCTCGCTGCTCTCTGGAGACCTTGGGTGGCT 2535
Cy 470 HisCysLysLeuProAspAlaValCysArgAspLeuSerGluAlaLeuArgAlaAlaPro 489
Db 2536 GGCTGTGGCTCTCACAGCTGAGCTGAAGGACCTTGCTTTGGGCTGAGAGCAACAG 2595
Cy 490 AlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeu 509
Db 2596 ACCGTGACCGAGCTGGACCTGAGCTTCAATGTGCTACGATGCTGGAGCAACACCTT 2655
Cy 510 SerGluGlyLeuAlaTrpProGlnCysArgValGlnThrValArgValGlnLeuProAsp 529
Db 2656 TGCCAGACCTGACACAGCCGAGCTGCAGCTACAG-----CGACTGCAGCTG----- 2703
Cy 530 ProGlnArgGlyLeu-----GlnTyrLeuValGlyMetLeuArgGlnSer 544
Db 2704 GTCAGCTGTGGCTCACCTCTGACTGCTGCTGCCAGGACCTGGCTCTGTGCTTAGTGCCAGC 2763
Cy 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
Db 2764 CCCAGCTGAAGGAGCTAGACCTGCAGCAGAAACCTGTGATGATGCTGGCGTGCAGCTG 2823
```

```
QY 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAlaSerVal 584
Db 2824 CTCGTGAGGGGCTCAGGCATCTCCCTGCAAACTCATACGCTGGGCTGGACCAGACA 2883
Cy 585 GluLeuSerGluGlnSerLeuGlnGluLeuGlnAlaValLysArgAlaLysProAspLeu 604
Db 2884 ACTCTGATGTATGATGATGAGGAGGAACCTGAGGGCCCTGGAGCAGGAAACCTCAGCTG 2943
Cy 605 ValIle 606
Db 2944 CTCATC 2949
RESULT 11
US-09-931-071-5
; Sequence 5, Application US/09931071
; Patent No. US200201282:9A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Alnemzi, Ebad S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-931-071-5
Alignment Scores:
Pred. No.: 1,456-76 Length: 4287
Score: 762.00 Matches: 220
Percent Similarity: 48.79% Conservative: 103
Best Local Similarity: 33.23% Mismatches: 207
Query Match: 23.68% Indels: 132
DB: 10 Gaps: 22
US-10-029-347-2 (1-625) x US-09-931-071-5 (1-4287)
QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuLeuAspGlyAlaAspGluLeuPro--- 19
Db 1192 ATCTGTCTAGCCAGAGCGGCTGCTCTTCA7CTCGATGGTGTAGATGAG---CCAGGA 1248
Cy 20 ---AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGly 38
Db 1249 TGGGTCTTGAGGAGCCGAGTCTCGAGCTCTGTCTGCACTGGAGCCAGCCAGCGGCG 1308
Cy 39 AlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuVal 58
Db 1309 GATGCACTGCTGGGCGATTGCTGGGGAATACTATCTTCCGAGGCATCTCTCTGATC 1368
Cy 59 ThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla 78
Db 1369 AGGGCTCGGACCCAGCTCTGCAGAACCTCATTCCTTTTGGAGCAGCGACGTTGGGTA 1428
Cy 79 GluValArgGlyPheSerAspLysAspLysLysTyrPheTyrLysPhePheArgAsp 98
Db 1429 GAGGTCTCTGGGTTCTCTGAGTCCAGCAGGAGGAGNATATTTCTACAGATATTTACAGAT 1488
Cy 99 GluArgArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeu 118
Db 1489 GAAAGCAAGCAATTAGAGCTTTAGGTGTCATAATCAAAACAAAGAGCTCTGGGCCCTG 1548
Cy 119 CysPheValProPheValCysTyrIleValCysThrValLeuArgGlnGlnLeuLeu 138
Db 1549 TGTCTGTGCTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1608
```



```
QY 139 GlyArgSerLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheLeuThr 158
DB 1609 AAGGAAAACTCACATGACTCTCCAGAGACACACACACCTCTGTCTACATACCTTGGC 1668
QY 159 SerValLeuSerSerAlaPrcValAlaAspGlyProArgLeuGlnGlyAspArgSer 178
DB 1669 CAGGCTCTCCAGCTCAGGCATTTG-----GGACCCAG-----CTCAGAGAC 1710
QY 179 LeuCysArgLeuAlaArgGluGlyValLeuGlyArgAlaGlnPheAlaGluLysGlu 198
DB 1711 CTCGTGCTCTGGCTCTGAGGCTCTGGCAAAAGAACCTTTTCAGTCCAGATGAC 1770
QY 199 LeuGluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysGlu 218
DB 1771 CTCAGGAAGCATGGTTAGATGGGGCATCATCTCCACC---TCTTTGAAGATG----- 1821
QY 219 LeuProGlyValLeuGluThrGlu-----ValThrTyrGlnPheLeuSer 235
DB 1822 -----GGTATTCTTCAAGAGCACCCCTCTGAGCTACAGCTTCATCACCCTGT 1875
QY 236 PheGlnGluPheLeuAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThr 255
DB 1876 TTCCAAAGATTCTTTCAGCAATGCTCTATGCTTGGAGAT----- 1917
QY 256 AlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisLeuVal 275
DB 1918 ---GAGAAGGGAGAGGT-----AAACATTCTAATTGCATC 1950
QY 276 Leu----- 277
DB 1951 ATAGATTTCGAAAGACGTAGACGATATGAATACATGGCTGTTCGGGCATCAACC 2010
QY 278 ThrArgPheLeuPheGlyLeuLeuSerAlaGluArgValArgAspIleGluArgHisPhe 297
DB 2011 ACACGTTCTCTATGGGCTCTTAACTGTATGAGTGTAGGGGAGAGAGATGGAAACATCTTT 2070
QY 298 GlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGln 317
DB 2071 CACTGCGGCTGTCTCAG-----GGGAGGAACCTGTATGACGTGGGTC----- 2112
QY 318 GlnGlyCysProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThr 337
DB 2113 -----CCGTCCCTGCAGCTGTGCTG 2130
QY 338 GluGluProGluGluGluGluGluGluProAsnTyrProLeuLeuLeuTyr 357
DB 2131 CTGCACCA-----CACTCTCGAGTCCCTCCAC 2160
QY 358 CysLeuTyrGluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGlu 377
DB 2161 TGCCTGTACGAGACCGGAACAAAGCTTCTTGACACAAAGTATGTCGCCCATTCGAAGAA 2220
QY 378 LeuAlaLeuGluLeuArgValArgPheCys-----ArgValAspValAlaValLeuSer 394
DB 2221 ATGGGCATG-----TGTGTAGAAACAGACATGGAGCTCTTATGTGCACT 2265
QY 395 TyrCysValArgCysCysPrcAlaGlyGlnAlaLeuArgLeuIleSerCysArg----- 412
DB 2266 TTCTGCATTAAATTCAGCGCCACGTGAGAGAGCTTCAGCTGATTGAGGGCAGGACGAC 2325
QY 412 ----- 412
DB 2326 AGATCAACATGGAGCCCAACCATGGTGTCTGTCAGTGGGTCCCACTACAGATGCC 2385
QY 413 -----LeuValAlaAlaGlnGluLysLysLysSerLeuGlyLysArgLeuGln 429
DB 2386 TATTTGGCAGATTCTCTCTTCCTCCGCTCCAAAGGTCAACGAGAAACCTG---AAGAGCTGGAC 2442
QY 430 AlaSerLeuGlyGlySerSerGlnGlyThrThrLysGlnLeuProAlaSerLeuLeu 449
DB 2443 CTAAAGT-----GGAACCTCG-----CTGAGCCACTCTCGCATG 2475
QY 450 HisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSerSerLeuThrLeuSer 469
```

```
DB 2476 AAGAGTCTTTTAAGACCCCTGAGACGCCCTCGCTGCCTCTCTGGAGACCCCTCGGTTGGCT 2535
QY 470 HisCysLysLeuProAspAlaValCysArgAspLeuSerGluAlaLeuArgAlaAlaPro 489
DB 2536 GGCTGTGGCTCTCACAGCTGAGGACTGCAAGGACCTTCCTTTGGGCTGAGAGCCCAACCAG 2595
QY 490 AlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeu 509
DB 2596 ACCGTGACCGAGCTGGACCTTGAGCTTCATGTGTCTACGATGCTGAGACCAACACCTT 2655
QY 510 SerGluGlyLeuAlaIleProGlnCysArgValGlnThrValArgValGlnLeuProAsp 529
DB 2656 TGCCAGAGACTGACACAGCCGAGCTGCAAGCTACAG-----CGACTGCAGCTG----- 2703
QY 530 ProGlnArgGlyLeu-----GlnTyrLeuValGlyMetLeuArgGlnSer 544
DB 2704 GTCAGCTGTGGCTTCACGCTCTGACTGTGCGCAGGACCTGCGCTGTAGTGCCAGC 2763
QY 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
DB 2764 CCCAGCTGAAGAGCTAGACCTGCACGACAGAACACCTGGATGACGTTGGCGTCGACTG 2823
QY 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAlaSerVal 584
DB 2824 CTCGTGAGGGGCTCAGGCATCTCTGCTGCAAACTCATACGCTGGGCTGGACACAGACA 2883
QY 585 GluLeuSerGluGlnSerLeuGlnGluLeuGlnAlaValLysArgAlaLysProAspLeu 604
DB 2884 ACTCTAGTGTATGATGATGAGGAGGAACTGAGGGCCCTGGAGCAGAGAAACCTCAGCTG 2943
QY 605 ValIle 606
DB 2944 CTCATC 2949
RESULT 12
US-09-388-221-1
; Sequence 1, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Reg
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4422)
US-09-388-221-1
Alignment Scores:
Pred. No.: 1,52e-76 Length: 4422
Score: 762.00 Matches: 220
Percent Similarity: 48.79% Conservative: 103
Best Local Similarity: 33.23% Mismatches: 207
Query Match: 23.68% Indels: 132
DB: 10 Gaps: 22
US-10-029-347-2 (-625) x US-09-388-221-1 (1-4422)
QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuPro--- 19
DB 1192 ATCTGTCTAGGCGAGCGGCTGCTCTTCATCTCTCGATGGTGTAGATGAG---CCAGGA 1248
QY 20 ---AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGly 38
DB 1249 TGGGTCTTTCAGGAGCGCGAGTCTGTAGCTCTGTGCACTGTGAGCCACGACCCGCGC 1308
```

QY 39 AlaArgValLeuGlyGlyLeuSerLysAlaLeuLeuPThrAlaLeuLeuVal 58
DB 1309 SATGCATGCTGGGCACTTTGCTGGGAAACATATCTCCCGAGGATCTTCCTGATC 1368
QY 59 ThrThrArgAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla 78
DB 1369 ACCGCTGGACCAACAGCTTCGAGAACCTCATCTCTTTGGAGCAGCAGCTGGGTAA 1428
QY 79 GluValArgGlyPheSerAspLysAspLysLysTyrPheTyrLysPheArgAsp 98
DB 1429 GAGGTCTCGGGTCTCTGAGTCCAGCAGGAAGATATTTACAGATATTTACAGAT 1488
QY 39 GluArgArgAlaGluArgAlaTyrArgPheValLysGluAsnGlnThrLeuPheAlaLeu 118
DB 1489 GAAAGCAAGCAATTAGAGCTTTAGGTGTTGTCATCAATCAACAAAGAGCTCTGGGCCG 1549
QY 119 CysPheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeu 138
DB 1549 TGTCTTGCTCCCTGGGTGCTTGCTGGCTGCCTTGGCTGATGAGCAGATGAAGCGG 1608
QY 139 GlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThr 158
DB 1609 AAGGAAAAGCTCACATGCACTTCCAGACCAACCAACCTCTGCTTACATACCTTGCC 1668
QY 159 SerValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsn 178
DB 1669 CAGGCTCTCAAGCTCAGCCATTG-----GGACCCCG-----CTCAGAGAC 1710
QY 179 LeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluCysGlu 198
DB 1711 CTCTGCTCTCTGGCTGCTGAGGGCATCTGGCAAAAAAGACCTTTTTCAGTCCAGATGAC 1770
QY 199 LeuGluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGlu 218
DB 1771 CTCAGSAGGATGGGTAGATGGGGCCATCATCTCCACC-----TCTTTGAGAGAG----- 1821
QY 219 LeuProGlyValLeuGluThrGlu-----ValThrTyrGlnPheIleAspGlnSer 235
DB 1822 -----GGTATCTCTCAAGACACCCCATCCCTCTGAGCTACAGCTTCATCTACCTCTGT 1875
QY 236 PheGlnGluPheLeuAlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThr 255
DB 1876 TTCCAAGAGTTCTTTGAGCAATGCTCTATGCTTGAGGAT----- 1917
QY 256 AlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuVal 275
DB 1918 ---GAGAGGGGAGAGGT-----AAACATTCTAATTGCATC 1950
QY 276 Leu----- 277
DB 1951 ATAGATTGGAAAAAGACGCTAGAGCATATGGAATACATGCGCTGTTGGGGCATCAACC 2010
QY 278 ThrArgPheLeuPheGlyLeuLeuSerAlaGluArgValArgAspIleGluArgHisPhe 297
DB 2011 ACAGCTTCTCTATGGGCTTTAAGTGATGAGGGGAGAGAGATGGAGAAACATCTTT 2070
QY 298 GlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlnGly 317
DB 2071 CACTGCGCGTGTCTCAG-----GGAGGAAACCTGATGAGTGGGT----- 2112
QY 318 GlnGlyCysProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThr 337
DB 2113 -----CCGTCCCTGCAGCTGCTG 2130
QY 338 GluGluProGluGluGluGluGluGluProAsnTyrProLeuGluLeuLeuTyr 357
DB 2131 CTGCAAGCA-----CACTCTGAGTCCCTCCAC 2160
QY 358 CysLeuTyrGluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGlu 377
DB 2161 TGCTTGTACGAGACTCGGAACAAAACGTTCTGTGACACAAAGTGATGGGCCCATTTGAAAGA 2220

QY 378 LeuAlaLeuGlnArgValArgPheCys-----ArgMetAspValAlaValLeuSer 394
DB 2221 ATGGGCATG-----TGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCACT 2265
QY 395 TyrCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuIleSerCysArg 412
DB 2266 TTCTGCAATTAATTCAGCGCCGACGCTGAAGAAGCTTCAGCTGATTGAGCGCAGGACAC 2325
QY 412 ----- 422
DB 2326 AGATCAACATGAGGACCCCAACCATGCTAGTCTGTTCAAGTGGTCCCAGTCACAGATCCC 2385
QY 413 -----LeuValAlaAlaGlnLulysLysLysSerLeuGlyLysArgLeuGln 429
DB 2386 TATTGGCAGATCTCTTCTCGCTCTCAAGGTCACCAAGAACCTG---AAGAGCTGCAC 2442
QY 430 AlaSerLeuGlyGlyGlySerSerGlnGlyThrLysGlnLeuProAlaSerLeuLeu 449
DB 2443 CTAAGT-----GGAAACTCG-----CTGAGCCACTCTGCAGTG 2475
QY 450 HisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSerSerLeuThrLeuSer 469
DB 2476 AAGAGTCTTTGTAAGACCTTGAGAGCCCTCGCTCGCTCTCGAGACCTCGGTTGGCT 2535
QY 470 HisCysLysLeuProAspAlaValCysArgAspLeuSerGluAlaLeuArgAlaAlaPro 489
DB 2536 GCGCTGGCCCTCACAGCTGAGGACTGCAAGGACCTTGCTTTGGGCTCAGAGCCAACAG 2595
QY 490 AlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeu 509
DB 2596 ACCCTGACCGAGCTGGACCTTCAATGTGCTCACGGATGCTGGAGCCAAACACCTT 2655
QY 510 SerGluGlyLeuAlaTrpProGlnCysArgValGlnThrValArgValGlnLeuProAsp 529
DB 2656 TCCAGAGACTGAGACGCCGAGCTGCAAGCTACAG-----CGACTGCAGCTG----- 2703
QY 530 ProGlnArgGlyLeu-----GlnTyrLeuValGlyMetLeuArgGlnSer 544
DB 2704 GTCACTGTGGGCTTCACGTCTGACTGTGCCAGGACCTGGCTCTGTCTTAGTSCCAGC 2763
QY 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
DB 2764 CCAGGCTGAAGAGAGCTAGACCTGCAGCAGAACCAACCTGGATGACGTTGGCTGCGACTG 2823
QY 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAlaSerVal 584
DB 2824 CTCTGTGAGGGCTCAGGCATCTCTCGTCAAACTCATACGCTGGGGCTGAGCAGACA 2883
QY 585 GluLeuSerGluGlnSerLeuGlnGluLeuAlaValLysArgAlaLysProAspLeu 604
DB 2884 ACTCTGAGTGATGAGATGAGGAGGAGAACTGAGGGGCTTGGAGCAGAGAAACCTCAGCTG 2943
QY 605 ValIle 606
DB 2944 CTCATC 2949
RESULT 13
US-09-388-221-9
; Sequence 9, Application US/09388221A
; Publication NO. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US 20020192643A1e1 Card Proteins Involved in Cell Death Re
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 4556
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

; NAME/KEY: CDS
 ; LOCATION: (1,1) (4362)
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: Construct
 US-09-388-221-9

Alignment Scores:
 Pred. No.: 1,598-76 Length: 4556
 Score: 762.00 Matches: 220
 Percent Similarity: 48.79% Conservative: 203
 Best Local Similarity: 33.23% Mismatches: 207
 Indels: 232
 Query Match: 132
 DB: 22

US-10-029-347-2 (1-625) x US-09-388-221-9 (1-4556)

```

Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuLeuAspGlyAlaAspGluLeuPro--- 19
Db 1192 ATCCCTGTCTAGCCACAGCGGCTGCTTCTCATCTCGATGGTGTAGATGAG---CCAGGA 1248
Qy 20 ---AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGly 38
Db 1249 TGGGTCTTCAGGACCGGAGTCTGTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGGCG 1308
Qy 39 AlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuVal 58
Db 1309 GATGCACTCTGGGCAGTCTTGTGGGAAACTATATACCTCCGAGGCATCTTCTCTGATC 1368
Qy 59 ThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla 78
Db 1369 ACGGCTCGGACACACAGCTCTGACAGAACCTCATCTCTTCTTGAGCAGCGACGTGGGTA 1428
Qy 79 GluValArgGlyPheSerAspLysLysLysLysLysLysLysLysLysLysLysLysLys 98
Db 1429 GAGGCTCTGGGCTTCTCTGAGTCCACGACGAGGAGGATATTTCTACAGATATTTACAGAT 1488
Qy 99 GluArgArgAlaGluArgAlaTyrArgPheValLysGluAspGluThrLeuPheAlaLeu 118
Db 1489 GAAAGCAAGCAATAGAGCCCTTGGTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTG 1548
Qy 119 CysPheValProPheValCysTrpLeuValCysThrValLeuArgGlnLeuGluLeu 138
Db 1549 TGTCTTGTGCTCTGGGTGCTGCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1608
Qy 139 GlyArgAspLeuSerArgThrSerLysThrThrThrSerValThrLeuLeuPheLeuThr 158
Db 1609 AAGGAAAAACTCACACTGACTTCCAAAGACACACACACACCTCTGTCTACATTACCTTGC 1668
Qy 159 SerValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsn 178
Db 1669 CAGGCTCTCCAAAGCTCAGCCATG-----GGACCCCA-----CTCAGAGAC 1710
Qy 179 LeuCysArgLeuAlaArgGluGlyValLeuGlyValArgAlaGlnPheAlaGluLysGlu 198
Db 1711 CTCTGCTCTCTGCTCTGAGGCGATCTGGCAAAAGACCCCTTTTCACTCCAGATGAC 1770
Qy 199 LeuGlnGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGlu 218
Db 1771 CTCAGGAAGCATGGGTAGATGGGGCCATCATCTCCACC---TTCTTGAAGATG----- 1821
Qy 219 LeuProGlyValLeuGluThrGlu-----ValThrTyGlnPheLeuAspGlnSer 235
Db 1822 -----GGTATCTTCAAGAGCACCCCATCCCTCTGAGCTACAGCTTTCATCACCCTCTGT 1875
Qy 236 PheGlnGluPheLeuAlaLeuSerTyLeuLeuGluAspGlyValProArgThr 255
Db 1876 TTCCAAGAGTCTTTCAGCAATGCTCTATGCTTTGGAGGAT----- 1917
Qy 256 AlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuVal 275
Db 1918 ---GAGAGGGGAGAGGT-----AACATTCTTAATTGTCATC 1953
Qy 276 Leu-----Thr 277
  
```

```

Db 1951 ATAGATTGGAAAAGACGCTAGAGCATATGGAATACATGGCCCTGTTTGGGGCATCAACC 2010
Qy 278 ThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAspLeuGluArgHisPhe 297
Db 2011 ACACGTTTCTCTATTGGGCTGTTTAAGTGTATGAGGGGAGAGAGATGGAGAACATCTTT 2070
Qy 298 GlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlnGly 317
Db 2071 CACTGCGCGCTGTCTCAG-----GGAGGAACCTGATGCAGTGGGTC----- 2112
Qy 318 GlnGlyCysProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThr 337
Db 2113 -----CCGTCCCTGCAGCTGCTG 2130
Qy 338 GluGluProGluGluGluGluGluGluProAsnTyrProLeuGluLeuLeuTyr 357
Db 2131 CTGACGCA-----CACTCTCTGGAGTCCCTCCAC 2160
Qy 358 CysLeuTyrGluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGlu 377
Db 2161 TGCTTGACGAGCTCGGAACAAACGTTCTCTGACACAGATGATGCCCATTTTCGAAGA 2220
Qy 378 LeuAlaLeuGlnArgValArgPheCys-----ArgMetAspValAlaValLeuSer 394
Db 2221 ATGGGCATG-----TGTGTAGAAACACAGATGGAGCTCTTAGTGTGCATC 2265
Qy 395 TyrCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuLeuSerCysArg 412
Db 2266 TTCTGCATTAAATTCAGCCGCCACGCTGAAGAGCTTCAGCTGATTGAGGCGCAGCAGCAC 2325
Qy 412 ----- 412
Db 2326 AGATCAACATGGAGGCCACCATGTTAGTCTCTTTCAGGTGGGTCCACAGTACACAGATGCC 2385
Qy 413 -----LeuValAlaAlaGlnGluLysLysLysLysLysLysLysLysLysLysLys 429
Db 2386 TATTGGCAGATTCTCTCTCGCTCAAGGTCACCAAGACCTG---AAGGAGCTGGAC 2442
Qy 430 AlaSerLeuGlyGlyGlySerSerGlnGlyThrThrLysGlnLeuProAlaSerLeuLeu 449
Db 2443 CTAAAT-----GGAACTCG-----CTGAGCCACTCTGCAGTG 2475
Qy 450 HisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSerSerLeuThrLeuSer 469
Db 2476 AAGAGTCTTTGTAAAGACCTGAGACGCCCTCGCTGCTCTGAGACCCCTGCGTGGCT 2535
Qy 470 HisCysLysLeuProAlaValCysArgAspLeuSerGluAlaLeuArgAlaAlaPro 489
Db 2536 GGCTGTGGCTCACAGCTGAGGACTGCAAGGACCTTGGCTTTGGGCTGAGAGCCAAACCG 2595
Qy 490 AlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeu 509
Db 2596 ACCCTGACCGAGCTGGACCTGAGCTTCAATGTCTCACGGATGCTGGAGCCAAACACCTT 2655
Qy 510 SerGluGlyLeuAlaTrpProGlnCysArgValGlnThrValArgValGlnLeuProAsp 529
Db 2656 TGGCAGAGACTGAGACAGCGAGCTGCAAGCTACAG-----CGACTGCAGCTG----- 2703
Qy 530 ProGlnArgGlyLeu-----GlnTyrLeuValGlyMetLeuArgGlnSer 544
Db 2704 GTCAGCTGTGGCTCAGCTGTGACTCTGCCAGGACCTGGCTCTGCTGTAGTGCACAGC 2763
Qy 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
Db 2764 CCAGCTGAGGAGCTAGACCTGACGACAAACACTGTGATGACCTGGCTGGCTGCCACTG 2823
Qy 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAlaSerVal 584
Db 2824 CTCTGTGAGGGCTCAGGCATCTGCTGCTGCAAACTCATACGCTGGGGCTGGACACAGACA 2883
Qy 585 GluLeuSerGluGlnSerLeuGlnGluAlaValLysArgAlaLysProAspLeu 604
  
```

```
Db 2884 ACTCTGAGTGATGAGTGGAGGAACTGAGGGCCCTGGAGCAGGAGAACTTCAGCTG 2943
Qy 605 Valile 606
Db 2944 CTCATC 2949

RESULT 14
US-09-996-617-1
; Sequence 1, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996, 617
; CURRENT FILING DATE: 2001-11-2
; PRIOR APPLICATION NUMBER: 09/931, 071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428, 252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340, 620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (523)...(4609)
US-09-996-617-1

Alignment Scores:
Pred. No.: 2,08e-76 Length: 5444
Score: 762.00 Matches: 220
Percent Similarity: 48.79% Conservative: 103
Best Local Similarity: 33.23% Mismatches: 207
Query Match: 23.68% Indels: 132
DB: 10 Gaps: 22

US-10-029-347-2 (1-625) x US-09-996-617-1 (1-5444)
Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeuPro--- 9
Db 1714 ATCTGTGTAGGCAGAGCGGTCTTTCATCTCGATGCTAGATGAG---CCAGGA 1770
Qy 20 ---AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGly 38
Db 1771 TGGGTCTTGAGGAGCCGAGTCTTGAGCTCTGTGCACTGGAGCCAGCCAGCCGCGG 1830
Qy 39 AlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuVal 58
Db 1831 GATGCACTGTGGCGACTTGTGTGGGAAACTATACTTCCGAGGCACTCTTCTCTGATC 1890
Qy 59 ThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla 78
Db 1891 ACGGCTCGGACCAAGCTCTGCAAGAACCTCATTCCTTCTTGAGCAGGCACTGGGTA 1950
Qy 79 GluValArgGlyPheSerAspLysAspLysLysLysTyrrPheTyrrLysPhePheAsp 98
Db 1951 GAGTGTCTGGGTTCTCTGAGTCCAGCAGGAGGAAATTTCTACAGATATTTACAGAT 2010
Qy 99 GluArgArgAlaGluArgAlaTyrrArgPheValLysGluAsnGluThrLeuPheAlaLeu 118
Db 2011 GAAAGCAAGCAATACAGCCTTTAGGTGGTCAAAACAAACAAAGAGCTCTGGGCCCTG 2070
Qy 119 CysPheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeu 138
Db 2071 TGTCTTGTGGCCTGGGTGTCCTGGCTGCTGCACTTGCCTGATGAGCAGAGTGAAGCGG 2130
Qy 139 GlyArgAspLeuSerAspThrSerLysThrThrThrSerValTyrrLeuLeuPheIleThr 158
```

```
Db 2131 AAGGAAAACTCACACTGATCTCCAGACCACCAACCTCTCTCTACATTACCTTGCC 2190
Qy 159 SerValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsn 178
Db 2191 CAGGCTCTCCAAGCTCAGCCATTG---GGACCCAG---CTCAGAGAC 2232
Qy 179 LeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGlu 198
Db 2233 CTCTGCTCTCTGGCTGTGAGGGCATCTGGCAAAAAGAACCCCTTTTCAGTCCAGATGAC 2292
Qy 199 LeuGluGlnLeuLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGlu 218
Db 2293 CTCAGGAAGCATGGGTAGATGGGGCCATCATCTCCACC---TTCTTGAAGATG--- 2343
Qy 219 LeuProGlyValLeuGluThrGlu-----ValThrTyrrGlnPheLeuAspGlnSer 235
Db 2344 -----GGTATTCTTCAAGACACCCCATCCCTCTGAGCTACAGCTTCATTCACCTCTGT 2397
Qy 236 PheGlnGluPheLeuAlaLeuSerTyrrLeuLeuGluAspGlyGlyValProArgThr 255
Db 2398 TTCCAAGAGATTCTTTGACGCAATGTCTATGCTTTGGAGGAT----- 2439
Qy 256 AlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuVal 275
Db 2440 ---GAGAAAGGGGAGAGGT-----AAACATTCTAATGTGATC 2472
Qy 276 Leu-----Thr 277
Db 2473 ATAGATTGGAAAAAGACCGTAGAAGCATATGGAATACATGCGCTGTTGGGGCATCAACC 2532
Qy 278 ThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPhe 297
Db 2533 ACACGTTTCTTATTGGGCTGTTAAGTGATGAGGGGAGAGAGATGAGGAACATCTTT 2592
Qy 298 GlyCysMetValSerGluArgValLysGlnAlaLeuArgTrpValGlnGlnGly 317
Db 2593 CACTGCCGCGCTGTCTCAG-----GGAGGAACCTGATGTCAGTGGTGC--- 2634
Qy 318 GlnGlyCysProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThr 337
Db 2635 -----CGTCCCTCGAGCTGCTG 2652
Qy 338 GluGluProGluGluGluGluGluGluProAsnTyrrProLeuGluLeuLeuTyrr 357
Db 2653 CTGCAGCCA-----CACTCTCTGAGTCCCTCCAC 2682
Qy 358 CysLeuTyrrGluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGlu 377
Db 2683 TGCTGTACGAGACTCGGAACAAACGTTCTCTGACACAAGTGTATGGCCCATTTCGAAGAA 2742
Qy 378 LeuAlaLeuGlnArgValArgPheCys-----ArgMetAspValAlaValLeuSer 394
Db 2743 ATGGGCACTG-----TGTTAGAAAACAGACATGGAGCTCTTAGTGTGCACT 2787
Qy 395 TyrrCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuIleSerCysArg----- 412
Db 2788 TTCTGCATTAATTCAGCCGCCACGTGAAGAGCTTCAGCTGATTTAGGCGGAGGAGCAGC 2847
Qy 412 ----- 412
Db 2848 AGATCAACATGAGGCCCCACCATCGTAGTCTGTTCAAGTGGTCCCACTCACAGATGCC 2907
Qy 413 -----LeuValAlaAlaGlnGluLysLysLysSerLeuGlyLysArgLeuGln 429
Db 2908 TATTGGCAGATCTCTCTCCGTCCTCAAGGTTCACAGAAACCTG---AAGGAGCTGGAC 2964
Qy 430 AlaSerLeuGlyGlySerSerGlnGlyThrThrLysGlnLeuProAlaSerLeuLeu 449
Db 2965 CTAAAGT-----GGAACACTCG-----CTGAGCCACTCTGCAGTG 2997
Qy 450 HisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSerSerLeuThrLeuSer 469
```

```
Db 2938 AAGAGCTCTTTGTAGACCTTGAGACGCCCTCGCTCCCTCGAGACGCCCTGGCTGGCT 3057
Qy 470 HisCysLysLeuProAspAlaValCysArgAspLeuSerGluAlaLeuArgAlaLeuPro 489
Db 3058 GGCTGTGGCTTCACAGCTGAGAGCTGCAAGGACCTTGGCTTGGGCTGAGAGCCACACG 3117
Qy 490 AlaLeuThrGluLeuGlyLeuLeuHisAsnArgSerGluAlaGlyLeuArgMetLeu 509
Db 3118 ACCCTGACCGAGCTGAGCTTCAATGTGCTCAGGATGCTGGAGCCAAACACCTT 3177
Qy 510 SerGluGlyLeuAlaLeuProGlnCysArgValGlnThrValArgValGlnLeuProAsp 529
Db 3178 TGCCAGAGACTGAGACGCCAGCTGCAAGCTACAG-----CGACTGACGCTG----- 3225
Qy 530 ProGlnArgGlyLeu-----GlnTyrLeuValGlyMetLeuArgGlnSer 544
Db 3226 GTCACTGTGGCTCTACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3285
Qy 545 ProAlaLeuThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
Db 3286 CCAGCTGAGGAGCTAGACCTGACGAGAGAACACCTGGATGACGTTGGCGTGGGACTG 3345
Qy 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAlaSerVal 584
Db 3346 CTCCTGTAGGGCTCAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3405
Qy 585 GluLeuSerGluGlnSerLeuGlnGlnLeuGlnAlaValLysArgAlaLysProAspLeu 604
Db 3406 ACTCTGAGTGATGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3465
Qy 605 ValLeu 606
Db 3466 CTCATC 1471
```

RESULT 15

```
US-09-931-071-1
; Sequence 11, Application US/09931071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TYPE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO:
; LENGTH: 5444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (523)...(4809)
US-09-931-071-1
```

```
Alignment Scores:
Pred. No.: 2,08e-76 Length: 5444
Score: 762.00 Matches: 220
Percent Similarity: 48.79% Conservative: 103
Best Local Similarity: 33.23% Mismatches: 207
Query Match: 23.68% Indels: 132
DB: 10 Gaps: 22
```

US-10-029-347-2 (1-625) x US-09-931-071-1 (1-5444)

```
Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuLeuAspGlyAlaAspGlnLeuPro--- 19
Db 1714 ATCTCTGTCTAGCCAGGCGGTCTCTTCATCTCGATGGTGTAGATGAG---CCAGGA 1773
```

```
Qy 20 ---AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGly 38
Db 1771 TGGGTCTTGGAGGCGGAGCTTCTGAGCTCTGTCTGCACTGGAGCCACACAGCCGCG 1830
Qy 39 AlaArgValLeuGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuVal 58
Db 1831 GATSCACTGTGGCGAGTTTGTGGGAAACATATATCTTCCGAGGATCTCTCTGATC 1830
Qy 59 ThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla 78
Db 1891 ACGCTCGGACCCACAGCTCTGCAGAACCTCATCTCTTTTGGAGCAGGCGACCTGGTA 1950
Qy 79 GluValArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLys 98
Db 1951 GAGGTCTGGGTCTCTGAGTCCAGCAGGAAGAATATTTCTACAGATATTTTACAGAT 2010
Qy 99 GluArgAlaGluArgAlaLysArgPheValLysGluAsnGluThrLeuPheAlaLeu 118
Db 2011 GAAAGGCAACCAATTAGAGCTTTAGCTTGTGTCAAATCAAAACAAAGAGCTCTGGG 2070
Qy 119 CysPheValProPheValCysTyrPheValCysThrValLeuArgGlnGlnLeu 138
Db 2071 TGTCTGTGGCTGGGTGCTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2130
Qy 139 GlyArgAspLeuSerArgThrSerLysThrThrSerValTyrLeuLeuPheLeuThr 158
Db 2131 AAGGAAATCTCACACTGACTTCCAGACCCACCAACCTCTGTCTACATTTACCTTGC 2190
Qy 159 SerValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArg 178
Db 2191 CAGGCTCTCCAAGCTCAGCCATTG-----GGACCCGAG-----CTCAGAGAC 2232
Qy 179 LeuCysArgLeuAlaArgGluGlyValLeuGlyArgAlaGlnPheAlaGluLysGlu 198
Db 2233 CTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2292
Qy 199 LeuGluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysGlu 218
Db 2293 CTCAGGAAGCATGGTTAGTGGGCCATCATCTCCACC---TCTTGAAGAAG----- 2343
Qy 219 LeuProGlyValLeuGluThrGlu-----ValThrTyrGlnPheLeuAspGlnSer 235
Db 2344 -----GGTATTCTTCAAGAGACCCCATCCCTCTGAGCTACAGCTTCATTCACCT 2397
Qy 236 PheGlnGluPheLeuAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThr 255
Db 2398 TTCAGAGTCTTTTGGCGAATGCTCTATGCTTGGAGAT----- 2439
Qy 256 AlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuVal 275
Db 2440 ---GAGAAGGGGAGAGGT-----AAACATTCTAATTGATC 2472
Qy 276 Leu-----Thr 277
Db 2473 ATAGATTGGAAAAGACGCTAGAACCATATGGAATACATGGCCCTGTTTGGGGCATCAACC 2532
Qy 278 ThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAspLeuGluArgHisPhe 297
Db 2533 ACAGTTTCTTATTTGGGCTGTTAAGTGATGATGGGGGAGAGAGAGAGAGAGAGAGAG 2592
Qy 298 GlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTyrValGlnGlnGly 317
Db 2593 CACTGCGGCTGCTCAG-----GGAGGAACCTGATGCGAGTGGGTC----- 2634
Qy 318 GlnGlyCysProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThr 337
Db 2635 -----CCGTCCTCTGACGCTGCTG 2652
Qy 338 GluGluProGluGluGluGluGluGluProAsnTyrProLeuGluLeuLeuTyr 357
Db 2653 CTCAGGCCA-----CACTCTCTGGAGTCCCTCCAC 2682
Qy 358 CysLeuTyrGluThrGlnGlnAspAlaPheValArgGlnAlaLeuCysArgPheProGlu 377
```

[illegible]

Search completed: October 3, 2003, 18:21:34
Job time : 459 secs

791	DB	TGCGCAGATGCTGGGCCAGTCGACAGCGGTGCTCTTCTCTGAGACGGCGCGACGAGC	850
127	QY	TGCGCGCGCTGGGGGGCCCCAGCGCGCCCTGACAGACCCCTTCGAGGGCGCGAGCG	186
851	DB	TGCGCGCGCTGGGGGGCCCCAGCGCGCCCTGACAGACCCCTTCGAGGGCGCGAGCG	910
187	QY	GCOCGGCGGTCTAGCGGGCTGCTGAGTAAAGCGGTGTGCCACCGGCCCTCTCTCTG	246
911	DB	GCSCGGGTGCTAGCGGGCTGCTGAGCAAGCGGTGTGCCACCGGCCCTCTCTCTG	970
247	QY	TGACACGCGCGCGCGCGCCCCCGGGAGGCTGCAGGGCGCGCTGTGTCCCCGAGTGG	306
971	DB	TGACACGCGCGCGCGCGCCCCCGGGAGGCTGCAGGGCGCGCTGTGTCCCCGAGTGG	1030
307	QY	CCGAGGTGCCGCGCTTCTCCACAAGGACAAAGAAAGTATTTCTCAAAGTTCTTCGGG	366
1031	DB	CCGAGGTGCCGCGCTTCTCCACAAGGACAAAGAAAGTATTTCTCAAAGTTCTTCGGG	1090
367	QY	ATGAGAGGAGCGCGAGCGCGCTACCGTTCGTGAGGAGACGAGACGCTTCCGCGC	426
1091	DB	ATGAGAGGAGCGCGAGCGCGCTACCGTTCGTGAGGAGACGAGACGCTTCCGCGC	1150
427	QY	TGTGCTTCGTGCGCTTCGTGTGCTGTGATCGCTGACCGGTGCGCCGACAGCTGAGC	486
1151	DB	TGTGCTTCGTGCGCTTCGTGTGCTGTGATCGCTGACCGGTGCGCCGACAGCTGAGC	1210
487	QY	TGCGTTCGGGACGTGTGGCGACGTCCAGACACACCGTCAAGTGTACCTGCTTTCATCA	546
1211	DB	TGCGTTCGGGACGTGTGGCGACGTCCAGACACACCGTCAAGTGTACCTGCTTTCATCA	1270
547	QY	CCAGCGTTCTGAGTCTGGTCCGCTAGCCACCGGGCCCGGTGCAGGGCGACCTCGCA	606
1271	DB	CCAGCGTTCTGAGTCTGGTCCGCTAGCCACCGGGCCCGGTGCAGGGCGACCTCGCA	1330
607	QY	ATCTGTCGCCCTGGCCCGGAGGGCGTCTCGGACGACCGGCGCACTTTCGCGAGAGG	666
1331	DB	ATCTGTCGCCCTGGCCCGGAGGGCGTCTCGGACGACCGGCGCACTTTCGCGAGAGG	1390
667	QY	AACTGAGCAACTGGAGCTTCGTGGCTCCAAAGTGACAGCGCTGTCTCAGCAAAAAGG	726
1391	DB	AACTGAGCAACTGGAGCTTCGTGGCTCCAAAGTGACAGCGCTGTCTCAGCAAAAAGG	1450
727	QY	AGTTCGCGCGCTCTGGACAGAGTCACTACCAAGTTTCATCGACACAGAGTTTCAGG	786
1451	DB	AGTTCGCGCGCTCTGGACAGAGTCACTACCAAGTTTCATCGACACAGAGTTTCAGG	1510
787	QY	AGTTCCTCGCGGACCTGCTACTCTGTGAGGACGCGGGGTGCCAGACACCGCGCTG	846
1511	DB	AGTTCCTCGCGGACCTGCTACTCTGTGTGAGACGCGGGGTGCCAGACACCGCGCTG	1570
847	QY	GCSCGTTGGGACACTCCTCGTGGGACGCCACGCGCACAGCCACTTGGTCTCAACA	906
1571	DB	GCSCGTTGGGACACTCCTCGTGGGACGCCACGCGCACAGCCACTTGGTCTCAACA	1630
907	QY	CGCGCTTCCTTCGAGTCTGAGCGGAGCGGATGCGCGCATCGAGCGGCATCTCG	966
1631	DB	CGCGCTTCCTTCGAGTCTGAGCGGAGCGGATGCGCGCATCGAGCGGCATCTCG	1690
967	QY	GCTGCATGGTTTCAGACCGGTGTAAGCAGGAGCGCCTCGCGTGGGTGCAGGGACAGGAC	1026
1691	DB	GCTGCATGGTTTCAGACCGGTGTAAGCAGGAGCGCCTCGCGTGGGTGCAGGGACAGGAC	1750
1027	QY	AGGCTGCCCGGAGTGGCACCGAGAGTGACCGAGGGGGCAAGGGCTCGAGGACACCG	1086
1751	DB	AGGCTGCCCGGAGTGGCACCGAGAGTGACCGAGGGGGCAAGGGCTCGAGGACACCG	1810
1087	QY	AAGAGCCAGAGGAGGAGGAGGAGGAGGAGGCCAACTACCACTCGAGTTGCTGTACT	1146
1811	DB	AAGAGCCAGAGGAGGAGGAGGAGGAGGAGGCCAACTACCACTCGAGTTGCTGTACT	1870
1147	QY	GCCTGTACGACGCGAGAGCGGTTGTGCGCCAAAGCCCTGTGCGGTTCCCGAGC	1206

1871	GCCTGTACGAGACGACGAGGACGCGTTTGTGCGCCAAGCCCTGTGTCCGGTTC	CGCGAGC	1930
1207	TGGCGCTGCAGCAGTGCCTTCTGCGCATGACGCTGGCTGTCTTGAGCTACT	CGCGTGA	1266
1931	TGGCGCTGCAGCAGTGCCTTCTGCGCATGACGCTGGCTGTCTTGAGCTACT	CGCGTGA	1993
1267	GGTGTGCCCTGTCTGGCAGGCACTCGGGCTGATCAGCTCGCATTTGCTTGC	TGCGCAGG	1326
1991	GGTGTGCCCTGTCTGGCAGGCACTCGGGCTGATCAGCTCGCATTTGCTTGC	TGCGCAGG	2050
1327	AGAGAGAGAGAGAGCCTGGGAGAGGCGCTCAGGCCACGCTGGGTGGCGG	CAGATTCTC	1386
2051	AGAAGAAGAAGAAGCCTGGGAGAGGCGCTCAGGCCACGCTGGGTGGCGG	CAGATTCTC	2110
1387	AAGGCACCAAAACAACTGCCAGCTCCCTTCTTCATCCACTCTTT	-----	1433
2111	TGGGACCCAACTGGCTCCAGAAGTAGCCCTTTCGACCAACCCTGCTGTG	CATCTGCCCA	2170
1434	-----	-----	1433
2171	CACCTCCACAGACCCTCGSCTCCTCCAGGCAAGGCTTTTGCACAGATTCT	TTTGAATA	2230
1434	-----	-----	1433
2231	TAGCTCCATTACGCCCTGCCCCAGGGCTTGGCACTCTGTTGAGAGGATGA	ATGTCACGG	2290
1434	-----	-----	1433
2291	TGTTGGCAGGGCTGGSCCTGGGACCCCAAGACCCCATGCAATGACTGAC	CCACTGTGCC	2350
1459	ATCTGAGCAGCCTCAGCTGCTCCCACTGCAAACTCCCTGACGGGCTCTG	CGCGAGACCTTT	1518
2351	ATCTGAGCAGCCTCAGCTGCTCCCACTGCAAACTCCCTGACGGGCTCTG	CGCGAGACCTTT	2410
1519	CTGAGGCCCTGAGGCGAGCCCCGCCACTGACGGAGCTGGSCCTCTCTCCA	CAACAGGCTCA	1578
2411	CTGAGGCCCTGAGGCGAGCCCCGCCACTGACGGAGCTGGSCCTCTCTCCA	CAACAGGCTCA	2470
1579	GTAGGCGGAGCTCGTATGCTGATGAGGGCTTACCTGGCGCAGTGCAGG	STGCAGA	1638
2471	GTAGGCGAGACTCGTATGCTGATGAGGGCTTACCTGGCGCAGTGCAGG	STGCAGA	2530
1639	CGGTCAGGTTACAGCTGCTGACCCACAGAGGGCTCCAGTACCTCGTGGG	TATGCTTC	1699
2531	CGGTCAGGTTACAGCTGCTGACCCACAGAGGGCTCCAGTACCTCGTGGG	TATGCTTC	2593
1699	GGCAGAGCCCGCCCTGACCAACCTGGATCTCAGCGSCTGCCAATCTCCG	CCCCCATGG	1759
2591	GGCAGAGCCCTGCCCTGACCACCTGGATCTCAGCGSCTGCCAATCTCCG	CCCCCATGG	2650
1759	TGACCTACCTGTGTGAGTCTCTGAGCACAGGGATCGGCGCTGCAGACCT	CAGTCTGG	1818
2651	TGACCTACCTGTGTGAGTCTCTGAGCACAGGGATCGGCGCTGCAGACCT	CAGTCTGG	2710
1819	CCTCTGTGGAGCTGAGCGAGCAGTCACTCAGGAGCTTCAGGCTGTGAAG	AGACAAAGC	1878
2711	CCTCTGTGGAGCTGAGCGAGCAGTCACTCAGGAGCTTCAGGCTGTGAAG	AGACAAAGC	2770
1879	CGGATCTGGTCACTCACACCCAGCGCTGAGCGGCCACCCACAACTCTCC	AAAGAACTCA	1938
2771	CGGATCTGGTCACTCACACCCAGCGCTGAGCGGCCACCCACAACTCTCC	AAAGAACTCA	2830
1939	TCTCGACCTTCTGA	1952	
2831	TCTCGACCTTCTGA	2844	

RESULT 2
US-10-066-521-17
; Sequence 17, Application US/10066521
; Publication No. US20030027757A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John

; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF

; FILE REFERENCE: 07334-341001
; CURRENT APPLICATION NUMBER: US/0/027,629
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 09/964,955
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/553,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506,067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6

; LENGTH: 3102

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-10-027-629-6

Query Match 8.4%; Score 226.4; DB 13; Length 3102;

Best Local Similarity 56.3%; Pred. No. 6.5e-46;

Matches 466; Conservative 0; Mismatches 356; Indels 6; Gaps 2;

Qy 14 GCGAGCCTGGCTGACCTGATCTGACACAGTCCCGACCGCGCGCGCGCGCGCGCA 73

Db 801 GAGGAGCCTGGGGACCTGATCATGAGCTGCTGCCCGACCCCAACCCACCCATCCAA 860

Qy 74 GATGCTGGCCAGCGCAGCGCTGCTTTCATCTGAGCGCGCGCGCGCGAGCTGC--C 130

Db 861 GATGCTGAGAAACCTCCAGAACTCTTCTCATGAGCGGTTTCGATGAGCTGCAAG 920

Qy 131 GCGCTGGGGGGCCCGAGCGCGCGCTGCACAGACCCCTTCGAGCGCGCGCGCGCGCG 190

Db 921 TGCCTTTGACGACACATAGACCGCTCTGACTGACTGACGACGAGCGCGGAGCA 980

Qy 191 GCGGTGCTAGGCGGCTGCTGAGTAAGCGCTGCTGCCACAGCGCCCTCTGCTGTGAC 250

Db 981 CATCTCTGAGCAGCGCTCATCAGAAAGAGCTGCTTCCGAGCGGCTCTCTCTCAT 1040

Qy 251 CACGCGCGCGCGCGCGCGCGGAGGTGAGGCGCGCGCTGTGTTCCCGCAGTGCACGA 310

Db 1041 CACGAGACCTGTGGCTTGAGAACTGCGACACTTGCTGGACCATCTCTGGCATGTGA 1100

Qy 311 GGTGCGCGGCTTCTCGACAGGACCAAGAAAGATATTTCTCAAGTCTTTCGSGATGA 370

Db 1101 GATCCTGGGTTCTCCGAGGCAAAAGAAAGAGTACTTCTTCAAGTACTCTCTGATGA 1160

Qy 371 GAGGAGCGCGCGCGCTACCGCTTCGTGAGGAGACGACGCTGTTCGCGCTGTG 430

Db 1161 GCGCAAGCGCGCGCGCGCTTTCAGTCTGATTCAGGAGACGAGGTCTCTTACCATGTG 1220

Qy 431 CTTCGTGCGCTTCTGTGCTGGATCGTGTGACCGCTGCTCGCGCAGCTGGAGCTCGG 490

Db 1221 CTTCATCCCGCTGCTGTGCTGATCGTGTGACTGAGCTGAAACAGACAGATGGAGTGG 1280

Qy 491 TCGGAGCTGTGCGGACGCTCAAGACCAACAGTCTAGTGTACTGCTTTTCTATACACAG 550

Db 1281 CAAGAGCTTTCGCCAGACATCTAAGACCAACCGCGGTGATGCTTCTTCTCTTCCAG 1340

Qy 551 CGTCTGAGCTGCGCTCCGTAGCGAGCGCGCGCTGCGAGCGAGCTGCGCAATCT 610

Db 1341 TTTCGTGACGCCCGGGAGCGGAGCAGGAGCAGCGCTCTGCGCCACCTCTGGGCGCT 1400

Qy 611 GTGCGGCTGGCCCGCGAGCGCGCTCTCGACCGCGCGCGAGCTTTGCCAAGAGAACT 670

Db 1401 CTGCTTTGGCTGCAGATGGAATCTGGAACCAAGAAATCTCTTTCAGGAGTCCGACCT 1460

Qy 671 GGAGCACTGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTTTCTCAGCAAAAGAGCT 730

Db 1461 CAGGAATCATGGACTGCGAAGGCGGAGTGT--GTCTGTCTTCTCAGGATGAACTGT 1517

Qy 731 GCGGGGCTGCTGAGACAGAGTCACTACAGTTCATCGACAGAGCTTCCAGAGTT 790

Db 1518 CCAAGAGAGTGGACTGCGAGAGTTCATAGCTTCATCCATGACTTTCAGGAGTT 1577

Qy 791 CTTGCGGCGACTGCTTACCTCTGAGGACGCGGGGTCCCGAGGAC 838
Db 1578 CTTTGGCGCATGTACTCTCTGGAAGAGAAAGGAAGGAGGAC 1625

RESULT 8

US-10-132-967-4

; Sequence 4, Application US/10/32967

; Publication No. US2003017084A1

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN PROTEIN FAMILY AND USES TH

; FILE REFERENCE: 07334-311001

; CURRENT APPLICATION NUMBER: US/10/132,967

; CURRENT FILING DATE: 2002-04-26

; PRIOR APPLICATION NUMBER: US/09/553,901

; PRIOR FILING DATE: 2000-09-01

; PRIOR APPLICATION NUMBER: US 09/506,067

; PRIOR FILING DATE: 2000-02-17

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 3857

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-132-967-4

Query Match 8.4%; Score 226.4; DB 12; Length 3857;

Best Local Similarity 56.3%; Pred. No. 6.9e-46;

Matches 466; Conservative 0; Mismatches 356; Indels 6; Gaps 2;

Qy 14 GCGCAGCCTGGCTGACCTGATCTCTGGACCACTGCCCCGACCGCGCGCGCGCGCGCA 73

Db 939 GAGGAGCCTGGGGACCTGATCATGAGCTGCTGCCCGACCCCAACCCACCCATCCAA 958

Qy 74 GATGCTGGCCAGCGCAGCGGCTGCTTTCATCTGAGCGCGCGCGAGCTGC--C 130

Db 999 GATCGTGAGAAACCCCTCCAGAACTCTTCTCATGACGGCTTCGATGAGCTGCAAG 1558

Qy 131 GCGCTGGGGGGCCCGAGCGCGCGCTGCACAGACCCCTTCGAGCGCGGAGCGCGCG 190

Db 1059 TGCCTTTGACGAGCACATAGGACCGCTCTGCACCTGACAGAGCGCGAGCGGAGGA 1118

Qy 191 GCGGTGCTAGCGGGCTGCTGAGTAAGCGCTGCTGCCACAGCGCTCTCTGCTGTGAC 250

Db 1119 CATCTCTGAGAGCTCTCATGAAAGAGCTGCTTCCGAGGCTCTCTGCTCATCAC 1178

Qy 251 CACGCGCGCGCGCGCGCGCGGAGCTGCGAGGCGCGCTGTGTTCCCGCAGTGCAGCGCA 310

Db 1179 CACGAGACCTGTGGCCCTGGAGAACTGCGACACTTCTGTSACCATCTCTCGGATGTGA 1238

Qy 311 GGTGCGGGCTTCTCCGACAGGACAGAGAGATATTTCTACAAGTCTTTCGGGATGA 370

Db 1239 GATCCTGGGTTTCTCCAGGCGCAAGAGAGAGTACTTCTTCAAGTACTTCTCTGATGA 1298

Qy 371 GAGGAGCGCGAGCGCGCTACCGCTTCGTGAAGGAGACGAGACGCTGTTCGCGGTGTG 430

Db 1299 GCGCCAGCAGGCGAGCGCTTCAGTCTGATTTCAGGAGACAGAGTCTCTTCCACATGTG 1358

Qy 431 CTTGCTGCGCTTCTGCTGATCGGTGTGACCGGTGCTGCGCGAGAGCTGAGAGCTCGG 490

Db 1359 CTTTCATCCCGCTGCTGCTGATCGGTGTGACCTGGACTGAAACAGCAGATGAGAGTGG 1418

Qy 491 TCGGAGCCTGTGCGGACGCTCCAGAGCCACAGTCTGCTGCTGCTGCTTTCATCACAG 550

Db 1419 CAAGAGCCTTGGCCAGACATCTAAGACCCACCGCGGTGTGCTTCTTCTTTCAG 1478

Qy 551 GGTTCCTGAGCTCGGCTCCGCTAGCGGACCGCGCGCTTTCAGAGCGGACCTTGGCGAATCT 610

Db 1479 TTTGCTGACGCCCGGGGAGGAGCGGAGCAGGCGCTCTGCGCCACCTCTGCGGGGCT 1538

Qy 611 GTGCGGCTGCGCGCGGAGGCGCTTCCTCGAGCGAGGCGGAGCTTTTCCGAGAGGAAGT 670


```
Db 1539 CTGCTCTTGGCTGACATGATCTGGAACACAGAAAATCTGTTTGAGGAGTCCGACCT 1598
Qy 671 GGAGCAACTGGAGTTGCTGCTCCTCAAGTGCAGACGCTCTTCTGAGCAAAAAGAGCT 730
Db 1599 CAGGAATCATGGACTGCAGAAAGCGGATGT---GTCGCTTTCTGAGGATGAACCTGTT 1655
Qy 731 GCCGGGCTGCTGGAGACAGAGGTCACTACCACTTCATCGACAGAGCTTCCAGGAGTT 730
Db 1656 CCAAAAGGAAGTGGACTGCGGAAGTTCTACAGCTTCATCTCACTGCTTTCCAGGAGTT 1715
Qy 791 CCTCGCGGCACTGCTTACCTCTGAGGAGCGCGGGTCCCGAGGAC 838
Db 1716 CTTTCCCGCATGTACTCTGCTGGAAGGAAAGGAAGGAGGAC 1763

RESULT 9
US-10-127-516-4
; Sequence 4, Application US/10127516
; Publication No. US20020187922A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Manji, Gulam A.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-370001
; CURRENT APPLICATION NUMBER: US/10/127,516
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 10/027,629
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 09/964,955
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506,067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (139)...(3240)
US-10-127-516-4

Query Match 8.4%; Score 226.4; DB 13; Length 3857;
Best Local Similarity 56.3%; Pred. No. 6.9e-46;
Matches 466; Conservative 0; Mismatches 356; Indels 6; Gaps 2;

Qy 14 GCGCAGCTGGCTCAGCTGATCTCGGACAGTGCAGCGCGCGCGCGCGCGCA 73
Db 939 GAGAGGCTGGGAGCTGATCATGAGTGTGCTGCCGACCCAAACCCACCATCCAA 998
Qy 74 GATGCTGCCGACGCGGAGCTGCTTTCATCTCGAGCGCGCGGAGCGAGTGC---C 130
Db 999 GATCGTGAGAAAACCCCTCAATCTCTCTCATGAGCGGCTTCGATGAGCTGCAAG 1058
Qy 131 GCGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 190
Db 1059 TGCCTTTGACGAGCACATAGACCGCTCTGCACTGACTGGCAGAAAGCGCGGGGAGA 1118
Qy 131 GCGGCTGCTAGCGGCTGCTGAGTAAGCGCTCTGCTCCACAGCGCCCTCTGCTGTGAC 250
Db 1119 CATCTCTGAGCAGCTTCAICAGAAAGAGCTGCTTCCGAGCGCTCTCTGCTCATCAC 1178
Qy 251 CACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 310
Db 1179 CACGAGACCTGTGGCCCTGGAGAACTGCAGCACTTGTGACCATCTCTCGGATGGA 1238
Qy 311 GGTGCGGCGCTTCTCGACAAAGGACAGAAAGATTTCTTACAGTTTCTTCCGGGATGA 370
```

```
Db 1239 GATCTCGGTTTCTCCGAGGCCAAAAGGAAGAGTACTTCTTCAAGTACTTCTCTGTATGA 1298
Qy 371 GAGGAGGCGCGAGCGCGCTTACCGTTTCGTGAAGAGAAACAGAGCGCTGTGGCGCTGTG 430
Db 1299 GCGCCAAAGCCAGGCGAGCTTTCAGTCTGATTCAGGAGAACGAGGTCTCTTTCACCATGTG 1358
Qy 431 CTTTCGTGCGCTTCTGCTGCTGATCGTGCACCGTGTGCGCCAGCAGCTTGGAGCTCGG 490
Db 1359 CTTTCATCCCGCTGCTGCTGATCGTGCACCTGCTGCACTGGACTGAAACAGCAGATGGAGTGG 1418
Qy 491 TCGGACCTGTGCGCCAGCTGCAAGACCAACGCTGAGTGTGCTGCTTTCATCACAG 550
Db 1419 CAAGAGCCTTGCACAGACATCTAAGACCAACCGCGGTGACGCTCTTCTCTTCCAG 1478
Qy 551 CTTTCGTGAGCTCGGCTCGGTAGCCGAGCGCGCGGTGCGAGGCGACCTGCGCAATCT 610
Db 1479 TTTGCTGAGCGCGCGGAGGAGCGAGGACGCGCTCTGCGCCACCTCTCTGGGGCT 1538
Qy 611 GTGCGCGCTGCGCGCGAGGCGCTCTCGGACGCGAGGCGGAGTTTGGCGAGAAAGAACT 670
Db 1539 CTGCTCTTTGGCTGCAGATGGAATCTGGAACAGAAATCTCTGTTGAGGAGTCCGACCT 1598
Qy 671 GGAGCAACTGAGCTTCTGGCTCCAAAGTGCAGCGCTGTTTCTCAGCAAAAAGAGCT 730
Db 1599 CAGGAATCATGACTGCAGAAAGCGGATGT---GTCTGCTTCTCTGAGGATGAACCTGTT 1655
Qy 731 GCGGCGCTGCTGGAAGACAGAGGTCACCTACAGTTCATCGACAGAGCTTCCAGGAGTT 790
Db 1656 CCAAAAGGAAGTGGACTGCGAGAGTTCTACAGCTTCATCCACATGACTTTCAGGAGTT 1715
Qy 791 CTTGCGGCACTGCTCTTACCTGCTGAGGAGCGCGGGGTGCCAGGAC 838
Db 1716 CTTGCGGCACTGTACTTACCTGCTGGAAGGAAAGGAAGGAGGAC 1763

RESULT 10
US-10-027-629-4
; Sequence 4, Application US/10027629
; Publication No. US20020197660A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Manji, Gulam A.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-341001
; CURRENT APPLICATION NUMBER: US/10/027,629
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 09/964,955
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506,067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (139)...(3240)
US-10-027-629-4

Query Match 8.4%; Score 226.4; DB 13; Length 3857;
Best Local Similarity 56.3%; Pred. No. 6.9e-46;
Matches 466; Conservative 0; Mismatches 356; Indels 6; Gaps 2;

Qy 14 GCGCAGCTGGCTGACCTGATCTCGGACAGTGCAGCGCGCGCGCGCGCGCA 73
Db 939 GAGAGGCTGGGAGCTGATCATGAGTGTGCTGCCGACCCAAACCCACCATCCAA 998
Qy 74 GATGCTGCCGACGCGGAGCTGCTTTCATCTCGAGCGCGCGGAGCGAGTGC---C 130
Db 999 GATCGTGAGAAAACCCCTCAATCTCTCTCATGAGCGGCTTCGATGAGCTGCAAG 1058
Qy 131 GCGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 190
Db 1059 TGCCTTTGACGAGCACATAGACCGCTCTGCACTGACTGGCAGAAAGCGCGGGGAGA 1118
Qy 131 GCGGCTGCTAGCGGCTGCTGAGTAAGCGCTCTGCTCCACAGCGCCCTCTGCTGTGAC 250
Db 1119 CATCTCTGAGCAGCTTCAICAGAAAGAGCTGCTTCCGAGCGCTCTCTGCTCATCAC 1178
Qy 251 CACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 310
Db 1179 CACGAGACCTGTGGCCCTGGAGAACTGCAGCACTTGTGACCATCTCTCGGATGGA 1238
Qy 311 GGTGCGGCGCTTCTCGACAAAGGACAGAAAGATTTCTTACAGTTTCTTCCGGGATGA 370
```


Db 999 GATCGTGAGAAAACCTCCAGAAATCTCTTCATGAGCGGTTTCGATGAGCTGCAAGG 1058
Qy 131 GCGGCTGGGGGGCCCCGAGGCGCGCTGTCAGACAGACCCCTTCGAGGCGCGAGCGGCG 190
Db 1059 TGCCTTTGACGAGCACATAGACCGCTCTGCACTGACTGCGAAGAGGCGGAGCGGAG 1118
Qy 191 GCGGGTGTAGCGGGCTGTGAGTAAGGCGCTGCTGCCACAGGCGCTCTCTGCTGTGAC 250
Db 1119 CATCTCTGAGGAGCTCAICAGAAAGAGTGTCTTCGAGGCGCTCTCTGCTATCAC 1178
Qy 251 CAGCGCGCGCGCGCGCGCGCGGCGTGAAGGCGCGCTGTGTTCGCCGAGTGGCGCGA 310
Db 1179 CACGAGACCTGTGGCGCTTGGAGAAATGCGAGCACTTGTGGACATCTCTCGGATGTGA 1238
Qy 311 GGTGGCGGCTTCTCCGACAAAGGACAGAGAGTATTTCTACAAGTCTTCGCGATGA 370
Db 1239 GATCTCTGGGTTTCTCGAGGCGAAAGGAGAGTCTTCTCAAGTACTCTCTGATGA 1298
Qy 371 GAGGAGGCGCGAGGCGCGCTTACCGCTTCTGTAAGGAGACGAGACGCTGTTCGCGCTG 430
Db 1298 GGCCTCAAGCCAGGCGAGCGCTTCACTCTGANTCAGGAGACGAGCTCTCTTCACTATG 1358
Qy 431 CTTGCGGCGCTTCTGCTGCTGATCGTGTGACCGTCTGCGGCGAGCTGAGCTCGG 490
Db 1359 CTTATCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1418
Qy 491 TCGGACCTGTGCGGCGAGCTCAAGACACACAGCTGAGTGTACTGCTTTTTCATCAC 550
Db 1419 CAAGAGCTTTCGCCAGACATCTAAGACACACAGCGGCTGCTGCTTCTTCTTCCAG 1478
Qy 551 GTTCTGAGCTCGGCTCGGCTAGCGAGCGGCGCGCTGCGAGGCGAGCTGCGGCT 610
Db 1479 TTTGCTGAGCGCGGCGGAGGAGCGAGGCGGCTTGTGCGGCGGCTTGTGCGGCG 1538
Qy 611 GTGCGGCTGCGCGCGGCGGCTTCTCGAGCGGCGGCGGCTTGTGCGGCGGAGGAG 670
Db 1539 CTGCTCTTGTGCTGAGTGGATCTGGAACACAGAAATCTTGTGAGGAGTCCGAGCT 1598
Qy 671 GGAGAACTGGAGCTTGTGCTGCTCAAGTGGAGAGCTGCTTCTCAGCAAAAGAGCT 730
Db 1599 CAGGAATCATGACTGCGAGAAAGCGGAGTGT---GTCTGCTTCTTCTGAGGATGA 1655
Qy 731 GCGGGGCTGTGAGACAGAGAGTCACTTACCTACAGTTCATCGACAGAGCTTCCAGAG 790
Db 1656 CAAAAGGAGTGGAGTGGCGGAGAGTCTACAGCTTCTTCTTCTTCTTCTTCTTCT 1715
Qy 791 CTTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 838
Db 1716 CTTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1763

RESULT 11

US-09-833-381-1976
; Sequence 1976, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1e1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1976
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1976

Query Match

7.5%; Score 200.4; DB 10; Length 1650;

Best Local Similarity: 55.0%; Pred. No. 1.5e-39;
Matches 437; Conservative 0; Mismatches 351; Indels 6; Gaps 2;
Qy 189 GCGCGGCTGCTAGGCGGCTGCTGAGTAAGCGCTGTGCTGCCACGCGCCTCTCTCTG 248
Db 247 GAGCTCTCTTAAACAGCTTAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCT 306
Qy 249 ACCACGCGCGCGCGCGCGCGCGGAGGCTGAGGCGCGCTGTTCGCCGAGGCGGC 308
Db 307 ACCACAGGCGCGCGCGCGCGCGCGGCTGAGGAGCTCCACCGCTGCTGGAGACCC 366
Qy 309 GAGGTGCGGCGCTTCTCGACAAAGCAAGAAAGTATTTCTACAAGTCTTTCGCGAT 368
Db 367 GAGATCTCTGGGCTTCTCTGAGGCAAGAAAGAACTACTTCTACAAGTATTTCCAA 426
Qy 369 GAGAGGAGGCGCGCGCGCGCTTACCGCTTCTGAGGAGGAGGAGAGCTGCTTTC 428
Db 427 GAGAGGAGGCGCGCGCGCGCGCGCGCTTACGTTACGAGGAGGAGGAGGCTTCT 486
Qy 429 TGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488
Db 487 TGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546
Qy 489 GGTGCGGAGCTTCTGCGGAGCTTCCAGACCGAGCTGCTGCTGCTGCTGCTGCTGCT 548
Db 547 GGGGGGCTGTTGAGACAGACGCTCCAGGACCACTGCTGCTGCTGCTGCTGCTGCTG 606
Qy 549 AGGCTTCTGAGCTCGGCTCCGCTAGCGAGCGGCGCGCGCTTGCAGGCGAGCTTGC 608
Db 607 AGTCTGATCAACCCCAAGCGCGGCGCGCGCTTCCAGCGCGCGCGCGCGCGCGG 666
Qy 609 CTGTCGCGCTGCGCGCGCGCGCGCTTCTCGGAGCGAGGCGGCTTTCGCGAGAGGAA 668
Db 667 TTTGCTCTTCTGCGGCGAGTGGGCTCTGGAATCAGAAATCTTATTTGAGGAGAG 726
Qy 669 CTGAGCAACTGAGCTTCTGCTGCTTCCAAAGTGCAGAGCTGTTTCTCAGCAAAAG 728
Db 727 CTCGGAAGACAGGCT---AGACGGGAGAGAGCTTCTGCGCTTCTCAACATGA 783
Qy 729 CTGCGGGGCTGCTGAGACAGAGCTTACCTACAGTTTCTGACGAGCTTTCAGGAG 788
Db 784 TTCCAGAGAGACATCACTGTGAGAGGAGTACAGCTTCTATCTTCTGAGTTTCCAG 843
Qy 789 TTCTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 848
Db 844 TTTCTTGGAGCTATGATATATCTGAGCGGGGAGGCGGCGGCGGCGGCGGCGGCG 903
Qy 849 GCGTTGGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 905
Db 904 GAGCTGAGCAGGCTGTTGACGAGTACGCTTTTCTGAAAGGAGCTTCTTCTGCACT 963
Qy 906 AGCGGCTTCTCTTCTGAGCTGCTGAGCGGAGCGGATCGCGAGCTGAGCGGCTT 965
Db 964 AGCGGCTTCTCTTCTGAGCTTCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 1023
Qy 966 GCGTGCATGCTTTC 979
Db 1024 TGCTGAAGTCTC 1037

RESULT 12

US-10-132-967-3
; Sequence 3, Application US/10132967
; Publication No. US20030170841A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07134-311001
; CURRENT APPLICATION NUMBER: US/10/132,967
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US/09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506,067

```
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3099
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3099)
US-10-132-967-3

      Query Match      6.7%; Score 179.2; DB 12; Length 3099;
      Best Local Similarity 54.0%; Pred. No. 2.9e-34;
      Matches 445; Conservative 0; Mismatches 358; Indels 21; Gaps 3;

QY 8 GTGGCGCGCAGCTGGTGTGATCTGATCTGGACCAAGTGCCTCCGACCGCGCGCGCGCGGT 67
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
675 GGGCCCGTGCAGTTTTCAGAGCTGGTCTTCAGGACCTGGCTGATTTGACAGCTGACAT 734
QY 68 GCGGAGATGCTGGCCAGCGCGCGCTGCTCTTCACTGACGCGCGCGCGCGAGCT 127
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
735 TCCACACATCTAGCCCAAGCAGCGGAAATCTTTCGTGATTCAGCGCTTGTAGCT 794
QY 128 G---CCGGCGCTGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 184
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
795 GGGAGCCGACCTGGGGCGCTGATCGAGGACATCTGCGGGGACTGGGAGAGAGAGAGCC 854
QY 185 CGGCGCGGGTGTAGGGGCGCTGCTGAGTAAGCGCTGCTGCCCGCGCGCGCGCGCTGCT 244
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
855 GGTGGCGTCCCTCCCTGGGGAGTTTCTGACAGGCTGATTTACCAAGCGCGCGCTGCT 914
QY 245 GGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTG 304
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
794 GGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTG 974
QY 305 GCGGAGGTGCGCGCTTCTCCGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 364
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
975 CATAGGGTGGAGGCTTCTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1034
QY 365 GCATGAGAGAGGCGCGCGCGCTACCGCTTCTGAGGAGAGAGAGAGAGAGAGAGAGAG 424
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1035 AGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1094
QY 425 GCTGTGCTTGTGCGCTTCTGCTGATCGTGTGACCGCTGCTGCCCGCGCGCGCGCGCTGGA 494
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1095 GCTGGGCTCGCGCGCGCGCGCGCTGCTGATCGTGTGACAGCTGTAAGCTGACAGTGA 1154
QY 485 GCTGGTTCGGGACCTGTCCGACAGCTTCAAGACAGAGAGAGAGAGAGAGAGAGAGAG 544
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1155 GAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1214
QY 545 CACACAGCTTCTGAGCTCGGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1215 CTGCAGC-----CGGTTCCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1259
QY 605 CAATCTGTCCGCGCTGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1260 GACGCTGAGGCTTCTGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1319
QY 665 GGAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1320 GGAATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1376
QY 725 GAGAGTCCGCGGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1377 CATCTCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1436
QY 785 GAGAGTCTCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1437 GCAGTGTCTCACTGCGCGCTTCTACACCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1480
```

D6	1260	GACCGT	GAGCTCT	CGCCGCG	CAGGSCCT	GTGGGCG	CAGAGCT	CGTGCT	TCACGAGA	1319
QY	665	GGAACT	GGAGCAACT	GGAGCTT	CGTGCGCT	TCCAAAGT	GTCAGACG	CTGTTT	TTCAGCAAAAA	724
DB	1320	GGATCT	GGAAGGCT	TCGGGGT	GCAGAGT	TCCGACCT	CC---G	CTGTTT	CCTGGACGAGA	1376
QY	725	GGAGCT	GC CGGCGT	GCTGTGG	GACAGAGT	CACTAC	CAAGTT	CATCTG	CACAGAGCTTCCA	784
DB	1377	CATCTC	CGCCAGAC	AGAGTCT	CCAAAGG	TGCTACT	CTCTTAT	CTCACC	TCAGCTTCCA	1436
QY	785	GGAGTT	CTCCGCGCA	CTGTCT	CACTGCT	GTGAGGAC	GC GCGG			825
DB	1437	SCAGTTT	CTCACTCCCT	GTGTTT	TACACCT	CTGGAGAGG	GAGGAGG			1480

RESULTS: 4

```

US-10-027-629-3
/ Sequence 3, Application US/10027629
/ Publication No. US20020197660A1
/ GENERAL INFORMATION:
/ APPLICANT: Bertin, John
/ APPLICANT: Maoli, Gallam A.
/ TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN
/ TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
/ FILE REFERENCE: 07J34-341001
/ CURRENT APPLICATION NUMBER: US/10/027,629
/ CURRENT FILING DATE: 2001-12-20
/ PRIOR APPLICATION NUMBER: US 09/964,955
/ PRIOR FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: US 09/653,901
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: US 09/606,067
/ PRIOR FILING DATE: 2000-02-17
/ NUMBER OF SEQ ID NOS: 17
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 3099
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-027-629-3

```

RESULT is

```

US-10-132-967-1
; Sequence 1, Application US/10132967
; Publication No. US20030170841A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-311001
; CURRENT APPLICATION NUMBER: US/10/132,967
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US/09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506,067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3431
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-967-1

```

RESULT is

```
Db      992  GGTCAACACCGCGCCAGGCGCTGAGGACCTCCGATCCTGCGGAGGAGCGGATCTA 1051
Qy      305  CGCCGAGGTGCGCGCTTCTCCGACAAAGAGAAAGTATTCTACAAAGTCTTCCG 364
Db      1052  CATAAGSGTGGAGGCTTCTGAGGAGGACAGAGGCGCTATTCTGAGACACTTGG 1111
Qy      365  GGATGAGAGGAGGCGGAGCGCGCTACCGCTTCTGTAAGGAGAAAGAGACGCTGTTCG 424
Db      1112  AGACGAGGACCAAGCCATGCTGCGCTTTGAGCTAATGAGGAGCAACGCGCGCTGTCCA 1172
Qy      425  GCTGTGCTTCTGCGCTTCGCTGATCGTGTGCACCGTGTGCGCCAGCAGCTGGA 484
Db      1172  GCTGGGCTCGGCCCGCGGCTGCTGATGCTGTGTCACCACTCTGAAGCTGCAGATGA 1231
Qy      485  GCTCGGTGGGACCTGTGCGGACGCTCCAAAGACCAACACGTCAGTGTACCTGCTTTCA 544
Db      1232  GAAGGGGAGGACCGGTCGCCACCTGCTCACCCGACGCGGGCTGTTCTGCGTTCTCT 1291
Qy      545  CACCAGGTTCTGAGCTGGGTCGCGTAGCGGAGCGGCGCGCGGTTGAGGCGACCTGCG 604
Db      1292  CTGCAGC-----CGGTTCCGCGAGGGCGCACAGCTGCGGGCGCGCTGCG 1336
Qy      605  CAATCTGTGCGCGCTGCGCCCGGAGGCGGCTCTCGGACGAGGCGCGAGTTTGCAGAA 664
Db      1337  GACGCTGAGCCTCTGCGCGCGGAGGCGCTGTGCGCGGAGACGTCGCTTCAACGAGA 1396
Qy      665  GGAACCTGAGAGCAACTGGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTTCTCAGCAAAA 724
Db      1397  GGA*CTGGAAAGGCTCGGG*GCAGGAGTCCGACCTCC--GTCTGT*CTTGGACGGAGA 1453
Qy      725  GGAGTCGCGGCGCTGCTGGAGACAGAGGTCACTACGATTGATCGACGAGGCTTCCA 784
Db      1454  CATCTTCGCGCAGACAGAG*CTCAAAGGCTGCTACTCTCTTCACTCACTCAGCTTCCA 1513
Qy      785  GGAGTTCCTCGCGGCACTGTCTACCTGCTGGAGGAGCGGGGG 828
Db      1514  GCAGTTTCTCACTGCCCTGTTCTACACCTGGAGAGGAGGAGG 1557
```

Search completed: October 3, 2003, 13:42:26
Job time : 649 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OX nucleic - nucleic search, using sw model

Run on: October 3, 2003, 03:53:26 ; Search time 156 Seconds
(without alignments)
7608.200 Million cell updates/sec

Title: US-10-029-347-1

Perfect score: 2689

Sequence: 1 cggacgctggcgagcagc.....aaaaaaaaaaaaaaaaaaaaa 2689

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCFUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	427.6	15.9	678	US-09-016-434-1011	Sequence 1011, Ap
2	236.6	8.8	257	US-09-016-434-208	Sequence 208, Ap
3	85.4	3.2	1371	US-08-910-731-1	Sequence 1, Appl
4	85.4	3.2	1371	US-08-910-731-7	Sequence 7, Appl
5	85.4	3.2	1371	US-08-910-731-5	Sequence 5, Appl
6	69.8	2.6	1386	US-08-910-731-5	Sequence 5, Appl
7	61.2	2.3	1470	US-09-099-041A-27	Sequence 27, Appl
8	61.2	2.3	1470	US-09-245-281-27	Sequence 27, Appl
9	61.2	2.3	1470	US-09-207-359B-27	Sequence 27, Appl
10	61.2	2.3	1470	US-09-340-620A-27	Sequence 27, Appl
11	61.2	2.3	2859	US-09-099-041A-9	Sequence 9, Appl
12	61.2	2.3	2859	US-09-245-281-9	Sequence 9, Appl
13	61.2	2.3	2859	US-09-207-359B-9	Sequence 9, Appl
14	61.2	2.3	2859	US-09-340-620A-9	Sequence 9, Appl
15	61.2	2.3	3080	US-09-099-041A-25	Sequence 25, Appl
16	61.2	2.3	3080	US-09-245-281-25	Sequence 25, Appl
17	61.2	2.3	3080	US-09-207-359B-25	Sequence 25, Appl
18	61.2	2.3	3080	US-09-340-620A-25	Sequence 25, Appl
19	61.2	2.3	3382	US-09-099-041A-7	Sequence 7, Appl
20	61.2	2.3	3382	US-09-245-281-7	Sequence 7, Appl
21	61.2	2.3	3382	US-09-207-359B-7	Sequence 7, Appl
22	61.2	2.3	3382	US-09-340-620A-7	Sequence 7, Appl
23	61.2	2.3	4302	US-09-245-281-38	Sequence 38, App
24	61.2	2.3	4302	US-09-207-359B-38	Sequence 38, App
25	61.2	2.3	4302	US-09-340-620A-38	Sequence 38, App
26	61.2	2.3	32042	US-09-245-281-44	Sequence 44, Appl
27	61.2	2.3	32042	US-09-340-620A-63	Sequence 63, Appl

28 56.6 2.1 1374 2 US-08-910-731-3 Sequence 3, Appl
29 56.6 2.1 1374 2 US-08-795-395-3 Sequence 3, Appl
30 55.2 2.1 1926 4 US-09-249-585A-2 Sequence 2, Appl
31 55.2 2.1 1926 4 US-09-410-393-3 Sequence 3, Appl
32 55.2 2.1 2580 3 US-09-050-863-2 Sequence 2, Appl
33 55.2 2.1 2580 4 US-09-359-081-2 Sequence 1, Appl
34 55.2 2.1 5452 2 US-09-130-114-1 Sequence 1, Appl
35 55.2 2.1 8705 4 US-09-647-344A-14 Sequence 14, Appl
36 55.2 2.1 9600 3 US-08-910-647-1 Sequence 1, Appl
37 55.2 2.1 9600 4 US-09-620-925-1 Sequence 15, Appl
38 55.2 2.1 10596 1 US-07-884-811-15 Sequence 15, Appl
39 55.2 2.1 10596 1 US-07-885-971-15 Sequence 15, Appl
40 55.2 2.1 10596 1 US-08-087-783A-15 Sequence 15, Appl
41 55.2 2.1 10596 2 US-08-194-088B-15 Sequence 15, Appl
42 55.2 2.1 10596 2 US-08-194-087-15 Sequence 15, Appl
43 55.2 2.1 10596 5 PCT-US93-04648-15 Sequence 15, Appl
44 53.8 2.0 690 4 US-09-252-991A-15944 Sequence 15944, A
45 53.8 2.0 1905 4 US-09-252-991A-16561 Sequence 16561, A

ALIGNMENTS

RESULT 1
US-09-016-434-1011
; Sequence 1011, Application US/C9016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1011:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 678 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927003
US-09-016-434-1011

Query Match 15.9%; Score 427.6; DB 4; Length 678;
Best Local Similarity 89.3%; Pred. No. 1.7e-73;

APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0342.3440003
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1371 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
NAME/KEY: CDS
LOCATION: 1...1368
US-08-910-731-1

Query Match 3.2%; Score 85.4; DB 2; Length 1371;
Best Local Similarity 54.7%; Fred. No. 7.7e-08;
Matches 192; Conservative 0; Mismatches 156; Indels 3; Gaps 1;
QY 1471 TCAGCGTGTCCCACTGCAAACTCCCTGACCGGTCTGCCGAGACCTTTCTGAGGCGCTGA 1530
DB 83 TCAGCGTGTCCCACTGCAAACTCCCTGACCGGTCTGCCGAGACCTTTCTGAGGCGCTGA 142
QY 1531 GGGCAGCCCCCGCACTGACGAGCTGGGCTCTCTCCACAAACAGGCTCAGTGGGCGGAC 1590
DB 143 GGGCAACCCCTCCCTGACCGAGCTGTGCTCCGACCAACGAGTGGGCGATGCCGCG 202
QY 1591 TGGGTATGCTGAGTGGGCTAGCTGCGCGAGTGCAGGCTGCAGAGCTCAGGCTAC 1650
DB 203 TGCACCTGGTGTGTCAGGCGCTGCAGAGCCCTGCAAGATCCAGAGCTCAGCTGC 262
QY 1651 AGTGTCTGACCC---CCAGCGAGGCTCCAGTACCTGTGGGTATGCTTCGGCAGAGCC 1707
DB 263 AGAAGTGTCTCCCTGACCGAGGCTGGGCTGCTGCCAGCAGCTGCGCTCCCTGC 322
QY 1708 CCGCCCTGACCACTCGATCTCAGGCGCTGCCACTGCCCGCCCTCATGTTGCTCGCAGAGCC 1767
DB 323 CCAGCTGCGGAGCTGCATCTCAGCGACCAACCACTGGGGGACGCGCGCTGCGGCTGC 382
QY 1768 TGTGTGAGTCTCTGACGACGAGGATGCGGCTGCAGACCTTCAGTCTGG 1818
DB 383 TCTGTGAGGCGCTCTCTGACCGCTCCAGTGCACCTGGAGAGCTGCAGTTGG 433

RESULT 4
US-08-910-731-7
Sequence 7, Application US/08910731
Patent No. 5932440
GENERAL INFORMATION:
APPLICANT: CHATTERJEE, DEB K.
APPLICANT: SHANDILYA, HARINI
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,731
FILING DATE: (Herewith)
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1371 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
US-08-910-731-7
Query Match 3.2%; Score 85.4; DB 2; Length 1371;
Best Local Similarity 54.7%; Fred. No. 7.7e-08;
Matches 192; Conservative 0; Mismatches 156; Indels 3; Gaps 1;
QY 1471 TCAGCGTGTCCCACTGCAAACTCCCTGACCGGTCTGCCGAGACCTTTCTGAGGCGCTGA 1530
DB 83 TCAGCGTGTCCCACTGCAAACTCCCTGACCGGTCTGCCGAGACCTTTCTGAGGCGCTGA 142
QY 1531 GGGCAGCCCCCGCACTGACGAGCTGGGCTCTCTCCACAAACAGGCTCAGTGGGCGGAC 1590
DB 143 GGGCAACCCCTCCCTGACCGAGCTGTGCTCCGACCAACGAGTGGGCGATGCCGCG 202
QY 1591 TGGGTATGCTGAGTGGGCTAGCTGCGCGAGTGCAGGCTGCAGAGCTCAGGCTAC 1650
DB 203 TGCACCTGGTGTGTCAGGCGCTGCAGAGCCCTGCAAGATCCAGAGCTCAGCTGC 262
QY 1651 AGTGTCTGACCC---CCAGCGAGGCTCCAGTACCTGTGGGTATGCTTCGGCAGAGCC 1707
DB 263 AGAAGTGTCTCCCTGACCGAGGCTGGGCTGCTGCCAGCAGCTGCGCTCCCTGC 322
QY 1708 CCGCCCTGACCACTCGATCTCAGGCGCTGCCACTGCCCGCCCTCATGTTGCTCGCAGAGCC 1767
DB 323 CCAGCTGCGGAGCTGCATCTCAGCGACCAACCACTGGGGGACGCGCGCTGCGGCTGC 382
QY 1768 TGTGTGAGTCTCTGACGACGAGGATGCGGCTGCAGACCTTCAGTCTGG 1818
DB 383 TCTGTGAGGCGCTCTCTGACCGCTCCAGTGCACCTGGAGAGCTGCAGTTGG 433

RESULT 5
US-08-795-395-1
Sequence 1, Application US/08795395
Patent No. 5965399
GENERAL INFORMATION:
APPLICANT: CHATTERJEE, DEB K.
APPLICANT: SHANDILYA, HARINI
TITLE OF INVENTION: Cloning and Expression of Rat Liver and
TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,395
FILING DATE: 04-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1371 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1368
US-08-795-395-1

Query Match 3.2%; Score 85.4; DB 2; Length 1371;
Best Local Similarity 54.7%; Pred. No. 7.7e-08;
Matches 192; Conservative 0; Mismatches 156; Indels 3; Gaps 1;
QY 1471 TCACCTGTCCCACTGCAATCTCCCTGACGGCTCTGCGAGACCTTTCTGAGGCCCTGA 1530
DB 93 TCAGCTCGACGACCTGCGCTCTACGGAGGACCTGCAAGACATCGTTCTGCCCTCC 142
QY 1531 GGGCAGCCCCCACTGACGGAGCTGGGCTCTCTCCACACAGAGCTCAGTACGCGGGAC 1590
DB 143 GGGCCCAACCCCTCCCTGACGAGCTCTGCTCCGACCAACGAGCTGGGGCATGCCGGC 202
QY 1591 TCGGTATGCTGAGTGAGGGCTAGCTGCGCGCTGCGAGGCTGAGAGGCTCAGGGTAC 1650
DB 203 TGCACCTGGTGTGAGGGGCTGCGAGAGCCCTGCAAGATCAGAACTCAGGCTGAC 262
QY 1651 AGCTGCTGACCC--CCACGACGAGGGCTCAGTACCTGGTGGTATGCTTCGCGAGAGCC 1707
DB 263 AGAAGCTGCTCCCTGACGAGGGGGCTGCGGGTCTCTCCACAGCAGCTGGCTCCCTGC 322
QY 1708 CGGCTGACCAACCTGGAATCTCAGGGGTGCACTGCGCCGCCCTGCTGACCTACC 1757
DB 323 CCACGCTGGGGAGTGCATCTCAGCGCAACCCCACTGGGGAGCCCGGCTGCGGTGC 382
QY 1768 TGTGTGACCTCTGACGACCAAGGATGCGGCTGCGAGCCCTCAGTCTGG 1813
DB 383 TCTGTGAGGGGCTCTCTGGAACCCCACTGACCTGCGAGAGCTGCGATTGG 433

RESULT 6
US-08-910-731-5
Sequence 5, Application US/08910731
Patent No. 5932440
GENERAL INFORMATION:
APPLICANT: CHATTERJEE, DEB K.
APPLICANT: SHANDILYA, HARINI

TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,731
FILING DATE: (Herewith)
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1386 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
US-08-910-731-5

Query Match 2.6%; Score 69.8; DB 2; Length 1386;
Best Local Similarity 51.9%; Pred. No. 7.5e-05;
Matches 205; Conservative 0; Mismatches 187; Indels 3; Gaps 2;
QY 1419 CTTATCCACTCTTTGAGGCAATGACTGACCCCACTGTGCACTCTGAGCAGCCTCAGCTG 1478
DB 388 CTGCACTGCTCTGCGAGGACTCTCTGGAGCCCCAGTGGCGCTGGAAGCTGCAGCTG 447
QY 1479 TCCACTGCAAACTCCCTGACGGGTCTGCCGAGACCTTTCTGAGGCCCTGAGGCAACC 1538
DB 448 GASTATTGAGGCTCTCGGCTCCAGCTGCGAGCCCTGGCTCTCAGGGCCCAAG 507
QY 1539 CCGCACTGACGAGCTGGGCTCTCTCCACAACAGGCTCAGTACGCGGAGCTGCTATG 1598
DB 508 CCGCACTTCAAGAGCTCAGGTTAGCAACAACAGACATCAATGAGGCTGGCTCGGTG 567
QY 1599 CTGAGTGAAGGCTTACGCTTGGCGGAGTGCAGGCTGACAGAGCTCAGG--GTACAGCTGC 1656
DB 568 CTGTGCAAGGCTGAAGGACTCTCCCTCCAGCTGGAGGCGCTCAAGCTGAGAGCTGC 627
QY 1657 CTGACCCCGAGGAGGCT--CCAGTACCTGGTGGTATGCTTGGCAGAGCCCCCGCTG 1715
DB 628 GGTGTGACATCAGACAACTGCGGGAGCTGTGCGCATTTGGCTCTCAAGGCCCTCGCTG 687
QY 1716 ACCACCTGATCTCAGCGGCTGCGAAGCTGCGCCCCCATGCTGACCTTACCTGTGTGCA 1775
DB 688 CGGAGCTGCGCTTGGGAGCAACAGCTGGGTGATGTGGGATGGCGAGCTGTGCCCA 747
QY 1776 GTCTTGCAGCACCAAGGATGCGGCTTGCAGAGCCCT 1810


```
QY 111 GACGGCGGAGAGCTCCCGGCTGGGGGCCCCGAGGCGGCGC---CTSCACAGAC 167
DB 733 GATGCTTGGAGAGCTGCACTGGAGCTTGACCTGAGCCGCTGCTGACAGCTCTGC 732
QY 168 CCGTTCGAGGGCGGAGCGGCGCGGGTGTAGCGGGTGTGAGTAAGGCGCTGCTG 227
DB 793 CCGTGGAGGCTTCCCGACCCCTGGCTTGTCTGGCCAACTGCTCAGTGGGAAGCTCTC 852
QY 228 CCCAGGCGCTTCTCTGCTGACACGCGCGCGCGCGCCCCCGGAGAGCTGCGAGGGCGC 287
DB 853 AAGGGGCTAGCAAGCTGCTACAGCCGSCACAGCATCGAGGTCCCGCGCCAGTTCCTG 912
QY 288 CTGCTTCCCGCAGTGGCGCGGAGGTGGCGGCTTCTCCGACAAGSACAGAAAGTAT 347
DB 913 CGGA-----AGAAAGTCTTCTCCGGGCTTCTCCCGACCCACTCGCGGCTAT 963
QY 348 TTCTACAAGTCTTCCGGGAAGAGAGGAGGCGCGGCTTACCGCTTGTGTAAGAGAG 407
DB 964 GCCAGGAGGATGTTCCCGCA3CGGGCCCTGCAGGACCGCTGCTGAGCCAGCTGGAGGCC 1023
QY 408 AACGAGAGCGCTGTTCCGCTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 461
DB 1024 AACCCCAACCTCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077

RESULT 10
US-09-340-620A-27
; Sequence 27, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-620A-27

Query Match 2.3%; Score 61.2; DB 4; Length 1470;
Best Local Similarity 50.5%; Pred. No. 0.0034;
Matches 209; Conservative 0; Mismatches 193; Indels 12; Gaps 2;

QY 51 GACCGGGCGCGCGGCTGGCGGAGATGCTGGCCAGCGCGAGCGGCTGCTTTCATCTG 1-0
DB 673 GACCCCGAGAGGATGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
QY 111 GACGGCGGCGAGCTGCTGGCGCTGGGGGCCCCGAGGCGGCGC---CTGCACAGAC 167
DB 733 GATGCTTGGAGAGCTGCACTGGAGCTTGAGACTTGGACCTGAGCCGCTGCTGAGAGCTCTGC 792
QY 168 CCGTTCGAGGGCGGAGCGGCGGGTGTAGCGGGTGTGCTGAGTAAGGCGCTGCTG 227
DB 793 CCGTGGAGGCTTCCCGACCCCTGGCTTGTCTGGCCAACTGCTCAGTGGGAAGCTGCTC 852
QY 228 CCCAGGCGCTTCTGCTGCTGACACGCGCGCGCGCGCCCCGAGGCGCTGCGAGGGCGC 287
DB 853 AAGGGGCTAGCAAGCTGCTACAGCCCGCACAGGCACTCGAGGTCCCGCGCCAGTTCCTG 912
QY 288 CTGCTTCCCGCAGTGGCGCTGAGGTGGCGGCTTCTCCGACAAGSACAGAAAGTAT 347
```

```
DB 913 CGGA-----AGAAAGTCTTCTCGGGGCTTCTCCCGACCCAGCTGCGCGCCTAT 963
QY 348 TTCTACAAGTCTTCTCCGGGAAGAGAGGCGCGGCTTACCGCTTGTGTAAGAGAG 407
DB 964 GCCAGGAGGATGTTCCCGCA3CGGGCCCTGCAGGACCGCTGCTGAGCCAGCTGGAGGCC 1023
QY 408 AACGAGAGCGCTGTTCCGCTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 461
DB 1024 AACCCCAACCTCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077

RESULT 11
US-09-099-041A-9
; Sequence 9, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2859
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-099-041A-9

Query Match 2.3%; Score 61.2; DB 4; Length 2859;
Best Local Similarity 50.5%; Pred. No. 0.0038;
Matches 209; Conservative 0; Mismatches 193; Indels 12; Gaps 2;

QY 51 GACCGGGCGCGCGGCTGGCGGAGATGCTGGCCAGCGCGAGCGGCTGCTTTCATCTG 110
DB 790 GACCCCGAGAGGATGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849
QY 111 GACGGCGGCGAGCTGCTGGCGCTGGGGGCCCCGAGGCGGCGC---CTGCACAGAC 167
DB 850 GATGCTTGGAGAGCTGCACTCGAGCTTGGACCTGAGCCGCTGCTGAGAGCTCTGC 909
QY 168 CCGTTCGAGGGCGGAGCGGCGGGTGTAGCGGGTGTGCTGAGTAAGGCGCTGCTG 227
DB 910 CCGTGGAGGCTGCGCCACCCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
QY 228 CCCAGGCGCTTCTGCTGCTGACACGCGCGCGCGCGCCCCGAGGCGCTGCGAGGGCGC 287
DB 970 AAGGGGCTAGCAAGCTGCTACAGCCCGCACAGGCACTCGAGGTCCCGCGCCAGTTCCTG 1029
QY 288 CTGCTTCCCGCAGTGGCGCGGAGGTGGCGGCTTCTCCGACAAGSACAGAAAGTAT 347
DB 1030 CGGA-----AGAAAGTCTTCTCGGGGCTTCTCCCGACCCAGCTGCGCGCCTAT 1080
QY 348 TTCTACAAGTCTTCCGGGAAGAGAGGCGCGGCTTACCGCTTGTGTAAGAGAG 407
DB 1081 GCCAGGAGGATGTTCCCGCA3CGGGCCCTGCAGGACCGCTGCTGAGCCAGCTGGAGGCC 1140
QY 408 AACGAGAGCGCTGTTCCGCTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 461
DB 1141 AACCCCAACCTCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1194

RESULT 12
US-09-245-281-9
; Sequence 9, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2003, 06:45:55 ; Search time 5518 Seconds
(without alignments)
11843.917 Million cell updates/sec

Title: US-10-029-347-1

Perfect score: 2689
Sequence: 1 cggacgcgtggcgcgagc.....aaaaaaaaaaaaaaaaaaaaa 2689

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 2278132 seqs, 1252238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1030.8	38.3	2730	11 AK039531	AK039531 Mus muscu
2	550.8	20.5	691	9 AI344276	AI344276 tc03ac7.x
3	411	15.3	427	9 AI762985	AI762985 w10sell.x
4	403.8	15.0	798	9 AW107005	AW107005 um1bd07.y

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	386	14.4	719	12	BG971110
6	339.4	12.6	603	12	BG971489
7	337.4	12.5	938	10	BF783992
8	316.4	11.8	670	10	BF633515
9	307.6	11.4	832	12	BG965959
10	268.8	10.0	416	13	BY262618
11	201.2	7.5	394	10	BF561521
12	200.6	7.4	377	9	AI558213
13	199.4	7.5	431	13	BY281505
14	198.8	7.4	532	9	AA541951
15	171	6.4	495	13	BX092006
16	167	6.2	3368	11	BC021272
17	167	6.2	3369	11	BC012789
18	157.8	5.9	409	9	AA997492
19	157.4	5.9	488	9	AI037137
20	152.8	5.7	921	13	BQ431649
21	138	5.1	664	29	AG067278
22	132	4.9	1110	13	BX457482
23	130.2	4.8	1156	10	BE731237
24	125.6	4.7	934	10	BE733220
25	122.8	4.6	629	14	CB016418
26	122.4	4.6	939	13	BX371966
27	120.8	4.5	124	14	C01661
28	113	4.2	440	28	AZ940081
29	111	3.8	577	28	AZ432279
30	101.2	3.7	646	10	BE265950
31	98.2	3.6	2394	11	AK028575
32	96.6	3.6	2520	11	AK032446
33	96.6	3.6	4476	11	AK028802
34	92	3.4	491	9	AI925881
35	89	3.3	506	14	CA556855
36	88.6	3.3	587	9	AV741792
37	86.2	3.2	3275	12	AK054426
38	85.4	3.2	568	12	B1344799
39	85.4	3.2	585	12	B1344410
40	85.4	3.1	589	12	BM030946
41	84.6	3.1	600	14	CA034391
42	84.6	3.1	595	9	AI883819
43	83.2	3.1	554	14	CB441719
44	83	3.1	253	14	CB606610
45	82.8	3.1			

ALIGNMENTS

RESULT 1
AK039531
LOCUS
DEFINITION
Mus musculus adult male spinal cord cDNA, RIKEN full-length enriched library, clone:A330055K17 product:similar to VASOPRESSIN RECEPTOR (Rattus norvegicus), full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

AK039531 2730 bp mRNA linear HRC 05-DEC-2002
Mus musculus adult male spinal cord cDNA, RIKEN full-length enriched library, clone:A330055K17 product:similar to VASOPRESSIN RECEPTOR (Rattus norvegicus), full insert sequence.
AK039531 GI:26087232
HTC: CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Ito, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

QY	784	AGGAGTTCTTCGGCGCACTGTCTTACTCTCTTGAGGAAGGCGGGTGCCCCAGGACGCCGG	843
DB	942	AGSAGTTCTTGGCCTGCATATTCTACTAGAAAGCTGAGCGAACCTCCGGGACGCCCCG	1001
QY	844	CTGGCGGGTTGGGACACTCTTCGGTGGGGAGCCCGAGCGCACAGCCACTTGGTGCTCA	953
DB	1002	CAGSCGGTGTSCAGAGCTCTTGAACTCTGAAGCGAGCTAGTGGTCTATCTTGCACTGA	1061
QY	904	CGACGGCTTCTCTTCGGAATCTCTGAGCGCGGAGCGGATCGCGACATCGAGCGCCACT	963
DB	1062	CAACCCGATTCTCTTTGGCTCTTAATAACAGAGGGGCTTCGTGACATTGGAAACCATT	1121
QY	964	TGCGCTGCATSGTTTCAGAGCTGTGAAGCAGAGAGCCCCTGGTGTGGTGCAGGGACGG	1023
DB	1122	TGGCTGTGTGGTCCAGATCATGTGAAGAAGACACCTCTGGGTGGT-----ACAG	1175
QY	1024	GACAGGGCTCCCCGGAGTGGCACCAAGAGGTGACCGAGGGGCGAAGAGGCTCGAGGACA	1083
DB	1176	GACAGAGCCACCCCAAGGGGCCCACTAGTAGGGCAAAAAGACTGCTGAGCTGGAGGACA	1235
QY	1084	CCGAAAGACGAGAGGAGGAGGAGGAGAGAG--CCCAACTACCTCTCTGAGGTTC	1140
DB	1236	TGAGGACGCGAGAGAGAGAGAGAGAGAGAGGAGGACTCAATTTGGACTGGAGTGC	1295
QY	1141	TGTACTGCTGTACGAGACGCGAGGAGACGGTTGTGGCCAAGCCCTGTGCGGTTCC	1200
DB	1296	TGTACTGCTGTATGAGACACAGAGGAGAGATTTGTTTCAGACGGCTCTCAGCAGCCCTC	1355
QY	1201	CGGAGCTGSCGCTCGACGGAGTGGCTCTTGCCGATGGAGGTGGCTCTCTGAGCTACT	1260
DB	1356	CAGAGATAGTACTGGACGGAGTCAGGTTGACCCGCAATGACCTTGAAGTCTTGAACTTACT	1415
QY	1261	CGTGTAGGTGTGCTCTGCTGTGAGGCACTCGGCTGATCAGCTGCAGATTTGGTTCTGTG	1320
DB	1416	GGTGTGAGTGTGCTCCAGATGTCAGGCTCTGAGACTGTGTGAGCTGTGGCTGGTGGCGG	1475
QY	1321	CGCAGGAGAAAGACAGAGAGCCTGGGNAAGCGGTCCAGGCCAGCTGGGTGGCGGCA	1380
DB	1476	CAAGGAGAGAGAGAGAGA-----AGAGAGCCTGTGAGCGGCTGAAGGGC-----	1525
QY	1381	GTCTCAAGGCAACACAAAAAACAATGCCAGCTCCCTTTCTTATCCACTCTTTCAAGCAA	1440
DB	1526	--TCTCAAGACCAAGAAACAACCCCCAGCTCTCCTTGTGTGGTCCACTCTGTGAGACAA	1583
QY	1441	TGACTGACCACTTGGCATCTGAGCAGGCTCAGCTGTCCACCTGCAAACTCCCTCAGG	1500
DB	1584	TGACTACCCCGAAATGTCACTGAGTGTTCATCTTTGCACCTGCAAGATCTCCCTGATG	1643
QY	1501	CGGTCTGCCGAGACCTTTCTGAGSCCCTGAGGCGAGCCCCCGACATGACGGAGTGGGCC	1560
DB	1644	CAGTTTGGCGAGACCTTTCGAGGSCCTGAAGGTAGCTCTCTCCCTAAGGAGTTGGGCC	1703
QY	1561	TCCTCCACACAGGCTCAGTGGCGGAGCTGGTATGTGTGAGTGAGGCGCTAGCCTGGC	1620
DB	1704	TCCTCCAGGCGGCTCACCAACAGAGGCTGGGTTTACTGTGCGAAGGCTTGCTTGGC	1763
QY	1621	CGCAGTGCAGGGTGCAGACGGTCAGGATACAGCTGCTGACCCCGCAGGAGGGTCTCAGT	1680
DB	1764	CCAAGTGCCAGGTGAAGACACTCAGGATGACGTGCTGACCTCCACAGAGGTGATCAATT	1823
QY	1681	ACCTGGTGGGTATGTTTGGGAGAGCCCCCTTGACCAACCTTGATCTCAGCGGCTGCC	1740
DB	1824	ACTTGGTCACTGCTGTCAAGCAGAGCCAGCTCTGACCACTCTGACCTCATAGTGGCTGTC	1883
QY	1741	AAC TGCCGCGCCCATGGTGACCTCACTGTGTGCAGTCTCTGCAGCACCGAGGATGGGCC	1800
DB	1884	AGCTGCTGGGCTCATTTGTGGACCTCTGTGTGCAGCCCTGAGCAGCCCTAAATGCTCCC	1943
QY	1801	TGCAGACCCCTCAGTCTGGCCTCTGTGAGCTGAGCGAGCAGTCACTACAGAGCTTCAGG	1860
DB	1944	TAAAGACCTCAGTCTGACCTCTGTGAGCTGAGTGAGAACTCACTGAGGATCTTCAAG	2003

Cy	186	CTGTGAAGAGCAAAAGCCGGATCTCTGGTATCATACACACCCAGCGCTGGACGGCACCCAC	1920
Db	2004	CTGTGAAGACATCAAGCCAGATCTGTCCATATATTCATATTCAAATTAAGGCACCTCACC	2063
Qy	1921	AACTCTCCCAAGAACTCATCTCGACCTTCTGTAGGCTCTGGTGGCCAGACGAGGCTGGAA	1983
Db	2064	GGCCTCGGAAG - GTGAGGCAGTCTCTTCTGAAGCTCCGAGTCCAGAACCATGAAG	2122
Qy	1981	ACCCT - - - AGTCAAGTCCCTGTGGAGAGACGGGCCATTCGAGGGCAGAGATATTG	2037
Db	2123	ATGCTGTAGGCCAAAGTCCCTCGAATGAACCTCCCATTCCAAAGCAAGAAAGATGCCA	2182
Qy	2038	CTCTCGGCTTTGGGAAACTTTTGAAGCCAGAGCGCCAGACAGGCACTGTGGAGGCCCA	2097
Db	2183	GCTTCTCCCTCCAGCAAGTCTTTCAAGCCACAGAAG - GGCAAGCAAAAGACCCAGTA	2240
Qy	2098	GAACGGACACCTGCCCGCTCCAGGACAGCCACAGACCTGSCCTCTCTCCACACCTGG	2157
Db	2241	GATACAGGCCCTTATAGCTCTCTGGATAGACCTGGCACGTGCGCTGCCCTGGACACTTG	2300
Qy	2158	SGTATCCCTTCTCCGCCAGGCCACCACTACTCCACCCACTTCTCT	2203
Db	2301	CGTTGTGCTTCTCCCTCAGTCCCAACACAGCAATGCTCTCTTCT	2349

```

RESULT 2
A1344276
LOCUS
DEFINITION
A:344276 691 bp mRNA linear EST 08-APR-1999
tctcaat7.x1 NCI CGAP Co-6 Homo sapiens cDNA clone IMAGE:2362740 3',
similar to TR:Q3035 Q63035 VASOPRESSIN RECEPTOR. ;, mRNA sequence.
ACCESSION
A1344276
VERSION
A1344276.1 GI:4081482
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 691)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP);
TITLE
Tumor Gene Index

```

Journal: J. Mol. Biol.

Comment: Contact: Robert Strausberg, Ph.D.
Email: cgapbs-irmail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.lnlnl.gov/brp/image/image.html
Insert Length: 1242 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 459.
Location/Qualifiers
1. 691
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="2062740"
/tissue_type="colon tumor, RER+ "
/lab_host="DH10B"
/clone_lib="NCI CGAP Col6"
/note="Organ: Colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Col6 was prepared, and ss circles were made in vitro. Following HAP hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloned IDs 1057416-1061255, and 1144584-1145351). Subtraction by Bento Soares and M. Fatima Bonafide. "

FEATURES

source


```

Db      599 TACCA 603

RESULT 7
BF783992
LOCUS      938 bp      mRNA      linear      EST 12-JAN-2001
DEFINITION 602107835F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4236272
5', mRNA sequence.
BF783992
ACCESSION  BF783992
VERSION     BF783992
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus

REFERENCE   1. (bases 1 to 938)
AUTHORS    N.H.-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-f@mail.nih.gov
           Tissue Procurement: Jeffrey E. Green, M.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM9844 row: e column: 09
           High quality sequence s-top: 660.
           Location/Qualifiers
             1..938
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="FVB/N"
               /db_xref="taxon:10090"
               /clone="IMAGE:4236272"
               /lab_host="DH10B (T1 phage-resistant)"
               /clone_lib="NCI CGAP Kid:4"
               /note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: Nori;
               Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
               Average insert size 1.75 kb. Constructed by Life
               Technologies. Note: this is a NCI_CGAP Library. "
BASE COUNT  262 a 232 c 24 g 170 t
ORIGIN
1
581  GCCCGCGTTTCAGGCGGACCTGCGCAATCTGTGCGCGCTGCGCGGAGGCGTCTCG- 639
|||||
358  ACCCGGGTTTCAGGCGAGCTGAGAACGCTGTGTGCTGCGCGGAGGAGGCGATCTCGA 417
|||||
640  --GAGCGCGGCGGCGATTTCGCGAGAGAACTGGAGCAACTGGAGCTTCTGCTGCTCAA 697
|||||
418  TCATCATTAAGCAACAATTTCTCAGAAAGAGGAGCTCGAGAAATTTGAAGCTTCGGGGTCCCA 477
|||||
698  AGTCAGACGCTGTTTCTCAGCAAAAAGGAGCTGCGCGGCGTGTCT-GGAGACAGAGGTCA 756
|||||
478  AGTTCAGACATATTTCTCAGACAGAGAGATACCAGAGTGTCTAAAAAAGTCA 537
|||||
757  CTACACAGTTTCATCAGCAGAGCTTCAGGAGTTCCTCGCGGACACTGTCCTACTCTGG 816
|||||
538  CTTACACAGTTTCATTGACAGAGCTTCAGGAGTTCCTGCTGCTGCTGCTGCTGCTAG 597
|||||
817  AGGAGCG-GCGGGGTGCCAGGACCGCGCTGCGCGCTTGGGACACTCTCTCGTGGGAC 875
|||||
598  AAGCTGAGCGGAATCTCCGCGGAGCGCCGACAGCGGTGTGCAGAAAGCTCTGAATCTTGAC 657
|||||
876  GCCCAGC 882
|||||
658  GCGGAAC 664

RESULT 8
BF633515
LOCUS      670 bp      mRNA      linear      EST 26-OCT-2001
DEFINITION BF633515 RIKEN full-length enriched, adult male spinal cord Mus
musculus cDNA clone A33055K17 5', mRNA sequence.
ACCESSION  BF633515
VERSION     BF633515
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus

REFERENCE   1. (bases 1 to 670)
AUTHORS    Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
           Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
           M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okada,
           Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
           D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
           Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
           Muramatsu, M. and Hayashizaki, Y.
           RIKEN Mouse ESTs (Arakawa, T., et al. 2002)
           Unpublished
           Contact: Yoshihide Hayashizaki
           Laboratory for Genome Exploration Research Group, RIKEN Genomic
           Sciences Center (GSC), Yokohama Institute
           The Institute of Physical and Chemical Research (RIKEN)
           1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0845, Japan
           Tel: 81-45-503-9222
           Fax: 81-45-503-9216
           Email: genome-res@gsc.riken.go.jp,
           URL: http://genome.gsc.riken.go.jp/
           Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
           M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
           Normalization and subtraction of cap-trapper-selected cDNAs to
           prepare full-length cDNA libraries for rapid discovery of new
           genes. Genome Res. 10 (10), 1617-1630 (2000)
           wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
           Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura,
           S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
           Hayashizaki, Y.
           RIKEN integrated sequence analysis (RISA) system--384-format
           sequencing pipeline with 384 multicapillary sequencer. Genome Res.
           10 (11), 1757-1771 (2000)
           Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
           Y. and Hayashizaki, Y.
           Computer-based methods for the mouse full-length cDNA
           encyclopedia: real-time sequence clustering for construction of a

```



```

Db 238 ACTGAGTCACTACAGTTCATTACACAGAGCTCCAGAGTCTTGCTGCTGCAATGCA 297
QY 807 TACCTGCTGAGGAGCGGGGTGCGGAGCGGCGGCTGCGGCGTGTGGACACTCTG 866
Db 298 TACCTACTAGAGCTGAGCGAAGTCTGGGAGCGCGCGGAGCGGTGTGCAGAACTCTG 357
QY 867 CGTGGGAGCGCCAGCGGACACGCACTTGTGTCTACACGCGGCTTCTCTTGGAGCT 926
Db 358 AACTCTGACGCGAGCTACGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 417
QY 927 CTGAGCGCGGAGCGGAGTGGCGACATCGAGCGGCTTGGCTGCTGCTGCTGCTGCTG 986
Db 418 CTAATATACAGAGGCGGCTTGTGACATTTGGAACCACTTTGCTGCTGCTGCTGCTGCT 477
QY 987 GTGAAGC-AGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1045
Db 478 GTGAACACAGGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531
QY 1046 ACAGAGGTGACGAGGCGGCTTAAGGCTTCAGAGACACCGAAGCGGAGGAGGAGGA 1205
Db 532 ACCAGTAGGCGGCAAAAGACTGCTGAGCTGAGGAGCATCGAGGACGAGGAGGAGGAGGA 591
QY 1136 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1162
Db 592 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 651
QY 1163 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1222
Db 652 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 710
QY 1223 GC-GCTTCTGCGCGCATGG-ACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1275
Db 711 CCAGGTGAGCGCGCATGGAACTGAGAGCTGGAATACGCGCATGCGAGTGGCC 766

```

RESULT 10

BY262618

LOCUS

DEFINITION

CDNA clone K33034H14 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

```

BY262618 416 bp mRNA linear EST 10-DEC-2002
BY262618 RIKEN full-length enriched, visual cortex Mus musculus
CDNA clone K33034H14 5', mRNA sequence.
BY262618.1 G1:26444130
EST.
Mus musculus (house mouse)
Xus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 416)
Nikaido, I., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Cojebori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beise, K. W., Blake, J. A., Bradt, D., Brusio, V., Chothia, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, J. J.,
Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltas, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G.,
Resole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reid, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wyrshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, I., Hara, A., Hashizume, M., Inotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata

```

K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Inotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numata, K., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissues were provided by Michela Fagioli and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hirotsawa, Wako-shi, Saitama 351-0199 Japan) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
 Location/Qualifiers
 1. 416
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strains="CS7BL/6J"
 /db_xref="taxon:10090"
 /clone="K33034H14"
 /tissue_type="visual cortex"
 /clone_lib="RIKEN full-length enriched, visual cortex"
 94 a 109 c 126 g 87 t

FEATURES

source

```

BASE COUNT 94 a 109 c 126 g 87 t
ORIGIN
Query Match 10.0%; Score 268.8; DB 13; Length 416;
Best Local Similarity 79.6%; Pred. No. 1.7e-16;
Matches 331; Conservative 0; Mismatches 82; Indels 3; Gaps 1;
QY 191 GCGGTGCTAGCGGGCTGCTGAGTAAGCGCTGCTGCCACGCGCTCTGCTGCTGCTGAC 250
Db 2 GAGAGTGTGAGCGGGCTGCTGAGCGAGGAGCTGTCAGGAGGAGTGTGCTGCGCAGTGGCAGA 61
QY 251 CACGGCGCGCGCGCGCGCGCGCGCTGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 310
Db 62 TACAGCGCATGCGCTGACGAGGAGGCTGCGAGGAGATGCTGCTGCGCAGTGGCAGA 121
QY 311 GGTGCGCGCTTCTCCGACAGGACAGAGAGTATTTCTCAAGTCTTCTCCGGGATGA 370
Db 122 AATACGCGCTTCTCAGACAAAGACAAAAGTATTTCTTCAAGTCTTCTCCGGGACGA 181
QY 371 GAGGAGCGCGCGCGCGCTACCGCTTGTGTAAGGAGACGAGCGCTGCTGCGCTGTG 430

```


Db 182 GAGGAGGACAGCGCCTACCGCTTCGTGAAGAGACGACGCTGTTGGACTGTG 241
 QY 431 CTTCGTGCTTCTGTCGTGATGTCACCGTCTGCGCCACGACGTGAGTCGG 490
 Db 242 CTTCGTGCTTCTGTCGTGATGTCACCGTCTGCGCCACGACGTGAGTCGG 301
 QY 491 TCGGACCTCTCGCGACGTCGAGACACGACGTCAGTGTACCTGCTTTTCATCCAG 550
 Db 302 CCGGACCTCTCTGTCACCTCCAGACCACTACTTCTGTGTAACCTCTTATCACCAG 361
 QY 551 CGTTCTGAGCTGCGCTCCGTAACCGACGCGGCCCGGTTTCAGGCGACGCTGCGCA 636
 Db 362 CATGCTCA--AGTCTCAGGACCAATGACCCCGGTTTCAGGAGAGCTGAGAA 414
 RESULT 11
 BF561521/c
 LOCUS
 DEFINITION
 UI-R-C0-hw-e-04-0-UI-1: 394 bp mRNA linear EST 12-DEC-2000
 UI-R-C0-hw-e-04-0-UI 5, mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (bases 1 to 394)
 AUTHORS
 Bernaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL
 GEORGE Res. 6 (9), 791-806 (1996)
 MEDLINE
 PUBMED
 97044477
 8889548
 COMMENT
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 CDNA Library Preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.resgen.com)
 This clone is also available through the I.M.A.G.E. Consortium at
 LNL (info@image.llnl.gov). IMAGE ID: 1781920
 Seq primer: M13 Forward.

FEATURES

source
 1. 394
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C0-hw-e-04-0-UI"
 /dev_stage="adult"
 /lab_host="DR10B (Life Technologies)"
 /clone_lib="UI-R-C0"
 /notes="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C0
 library is a subtracted library derived from the UI-R-A:
 and UI-R-E1 libraries. The UI-R-A1 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, and muscle. The UI-R-E1
 library consisted of a mixture of individually tagged
 normalized libraries constructed from 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dn track which
 allows identification of the library of origin of a clone
 within the mixture. The subtracted library (UI-R-C0) was
 constructed as follows: PCR amplified cDNA inserts from a
 pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had
 been derived was used as a driver in a hybridization with:

the pooled UI-R-A1 and UI-R-E1 library in the form of
 single-stranded circles. The remaining single-stranded
 circles (subtracted library) was purified by
 hydroxyapatite column chromatography, converted to
 double-stranded circles and electroporated into DH10B
 bacteria (Life Technologies) to generate the UI-R-C0
 library. This procedure has been previously described
 (Bonaldi, Lennon and Soares, Genome Research 6: 791-806,
 1996)."

BASE COUNT 87 a 117 c 93 g 97 t
 ORIGIN
 Query Match 7.5%; Score 201.2; DB 10; Length 394;
 Identical Similarity 72.6%; Pred. No. 3.4e-10;
 Matches 260; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
 QY 670 TSGAGCAACTGAGCTTCTGGCTCCAAAGTCGAGACGCTGTTTCTCAGCAAAAGAGCG 729
 Db 394 TGGAGAGATTGAAGCTTCAGGGTTCCCAAGTTCAGACAATGTTTCTCAGCAAGAGAGCG 335
 QY 730 TCCCGGGCGTCTCGAGACAGAGTCACTACAGTTTCATCGACCAAGAGCTTCCAGGAGT 789
 Db 334 TCCAGGAGAGTCTAGAAACTGTGTCACCTACCAGTTCATTGACCAAGAGCTTCCAGGAGT 275
 QY 790 TCCTCGGGGCACTCTCTACCTGCTGAGGACGGCGGTGCCAGGACCGCGCTGCGC 849
 Db 274 TCTTGGCTGCAATGTCATACCTACTAGCGCTGAGGAGGCCCGAGGAACCTCCGACGAA 215
 QY 850 CGCTTGGGACACTCCTCGTGGGACCCCGCGCACAGCCACTTGTGCTCACCACGC 909
 Db 214 GTGTGCAGATGCTCTGAACCTCTGACGCGGGCTGCGTGGTCACTGCGACTCACCACCTA 155
 QY 910 GCTTCCTCTTCGACTGCTGAGCGCGGAGCGGATCCCGACATCGAGCGGCACTTCGGCT 965
 Db 154 GATTCTCTTTTGGACTCTTAAGTACAGAGAGGATTCGTGACATTTGGAAACCATTTTGGCT 95
 QY 970 GCATGTTTTCAGAGCGTGTGAAGCAGGAGCGCCCTCGGTGGTGCAGGACAGCGGACA 1027
 Db 94 GTGTGTTGCCAGGGCGGTGTGNAACAGACACACTTTCGGTGGGTACAGGACAAAGGCA 37

RESULT 12

AI558213 377 bp mRNA linear EST 24-MAR-1999
 vj37a07.y1 Stratagene mouse diaphragm (#937303) Mus musculus CDNA
 clone IMAGE:931188 5' similar to TR:Q63035 Q63035 VASOPRESSIN
 RECEPTOR. ;, mRNA sequence.

ACCESSION

AI558213

VERSION

AI558213.1

KEYWORDS

GI:4508451

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 377)

AUTHORS

Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wy'et,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
 B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter
 E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.

TITLE

The WashU-NCI Mouse EST Project 1999

JOURNAL

Unpublished

COMMENT

Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MG:536108
 This read is a RESSEQUENCE of a previously sequenced mouse clone
 This read has been verified (found to hit its original self in the

correct orientation)

Possible reversed clone: similarity on wrong strand

Seq primer: -43RP from Gibco

High quality sequence stop: 369.

FEATURES

source
 1. 377
 /organism="Mus musculus"
 /mol_type="rRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:931198"
 /tissue_type="diaphragm"
 /dev_stage="adult"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Stratagene mouse diaphragm (#937303)"
 /note="Organ: diaphragm; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally from mRNA prepared from diaphragm muscle. Primer: Oligo dT. Average insert size: 1.5 kb. Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3."

BASE COUNT 88 a 114 c 93 g 82 t

Query Match 7.5%; Score 200.6; DB 9; Length 377;
 Best Local Similarity 70.9%; P-ed. No. 4e-10;
 Matches 266; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 1558 SGGCTCTCCACACAGGCTCATGTAGGGGGGACTCGTATCTGAGTGGAGGCTAGCCT 1617
 DB : GGTCTCTCAGAGCGGCTCACACACAGGCGCTTCTTACTGTGGAAGGCGCTGCTT 60
 QY 1618 GCGCGCAGTGCAGGGTGCAGAGCGTTCAGGTTACAGCTGCTCCGACCCAGGCGGCTCC 1677
 DB 61 GAGCCGAGTGCAGGTGAAGACTCAGAGTGCAGCTGCTTGGACTCCAGAGTGTATCA 120
 QY 1678 AGTACTGTGGGTATGCTTCGCGAGAGCCCGCCCTGACCCAGCTGATCTGAGCGGCT 1737
 DB 121 ATTACTTGGTATCTGCTGCTCAGAGAGCCAGTCTGACCACTCTGACCTCAGTGGCT 160
 QY 1738 GCGACTGCGCCCGCCATGTTGACTACCTGTCGAGTCTGTCAGCAGCAGGAGTGGG 1797
 DB 181 GTGAGTCTGCTGGGTCTATTTGGAACTCTTGTGAGCGCTGAGACCCCTAAAAGCT 240
 QY 1798 GCTCGAGACCTCAGTCTGGCTCTGTGGAGTGCAGCGAGCAGTCACTACAGGAGCTTC 1857
 DB 241 CCGTAAGACCTCAGTCTGACTCTGTGGAGTGCAGTGAAGTCACTTTCAGGAGCTTC 320
 QY 1858 AGGCTGTGAGAGCAAGCGGAGTGGTGTATCATACACACCCAGCGCTGACGCGCAC 1917
 DB 301 AGCTGTGAGAGCATCAAGCGGAGTCTGTCCATCATATATTCAAAATAGGCGACCTC 360
 QY 1918 CACAACCTCCCAAGG 1932
 DB 361 ACCGGCTCGGAAGG 375

RESULT 13

BY281505

LOCUS

DEFINITION
 cDNA clone K430354M18 5', mRNA sequence.

ACCESSION

BY281505

VERSION

BY281505.1 G1:26471842

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 431)

REFERENCE

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaide, I., Osato, N., Sato, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,

Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani,
 L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
 Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,
 P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perce, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,
 B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempie, C.A., Setou,
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,
 R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kogawa, I., Miyazaki, A., Sakai, K., Sasakawa, R., Shiba,
 K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,

JRC:http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane

T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.

, Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.

, Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami

X., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct

Submission.

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Michela Fagioli and Takao K. Hensch (

Laboratory for Neuronal Circuit Development Brain Science Institute

RIKEN 2-1 Hirotsawa, Wako-shi, Saitama 351-0198 Japan) whose

assistance we gratefully acknowledge. Please visit our web site

(http://genome.gsc.riken.go.jp) for further details.

FEATURES

Location/Qualifiers

1. 431

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="K430354M18"

/tissue_type="visual cortex"

/clone_lib="RIKEN full-length enriched, visual cortex"

109 a 113 c 114 g 95 t

BASE COUNT

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 58C, D-69120 Heidelberg, Germany
RZPD; :WAGp998M175911.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPEC3 (RZPDLB No.972)
<http://www.rzpd.de/Cloncards/cgi-bin/showLib.pl.cgi/response?libNo=972> Contact: Ina Rofls
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de

This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGAAACAGCTATGAC.

FEATURES

source

Location/Qualifiers
1..495
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998M175911 ; IMAGE:2380600"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: kidney; Vector: p7713-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco R1;
Plasmid DNA from the normalized library NCI CGAP Kid1 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323176-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 109 a 141 c 138 s 107 t
ORIGIN

Query Match 6.4%; Score 171; DB 13; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2480 GAGACTAGGCGAGGTGGAGACAGCGGAGAACCCCGCTGCTGGTGGGAGGATGACACAC 2539
Db 1 GAGACTAGGCGAGGTGGAGACAGCGGAGAACCCCGCTGCTGGTGGGAGGATGACACAC 50
Qy 2540 TGGTGGTGAGCGCCCGATGACGCTGACGCTAAATTCCTGCTGGACTCATTCCTGTG 2599
Db 61 TGGTGGTGAGCGCCCGATGACGCTGACGCTAAATTCCTGCTGGACTCATTCCTGTG 120
Qy 2600 GTTCTATTACCTGGCGAGCGGTGTACATAACAGGTCGGTGTCTACAA 2650
Db 121 GTTCTATTACCTGGCGAGCGGTGTACATAACAGGTCGGTGTCTACAA 171

Search completed: October 3, 2003, 13:31:35
Job time : 5527 secs

JOURNAL	Patent: WO 02061086-A 1 08-AUG-2002;	Db	601	TGCGCAATCTGTGCGGCTTGGCCCGCGAGGGGGTCTCTCGGAGCGAGGGCGCAGTTTGGCG	660
FEATURES	Bristol-Myers Squibb Company (US)	QY	661	AGAAGGAATGAGGACAACTGGAGCTTCTGTGGCTCTCAAAGTGACAGCCGTGTTTCTCAGCA	720
source	Location/Qualifiers	Db	662	AGAAGGAATGAGGACAACTGGAGCTTCTGTGGCTCTCAAAGTGACAGCCGTGTTTCTCAGCA	720
	1. .2689	QY	721	AAAAGGAGCTGCCGGGGTCTGGAGACAGAGGTCACTACCACTTCTGACACAGAGCT	780
CDS	/organism="Homo sapiens"	Db	721	AAAAGGAGCTGCCGGGGTCTGGAGACAGAGGTCACTACCACTTCTGACACAGAGCT	780
	/mol_type="genomic DNA"	QY	781	TCCAGGAGTTCTCTCGCGCACTGTCTACCTGCTGGAGGAGCGGGGTGCCCCAGACCG	840
	/db_xref="taxon:9606"	Db	781	TCCAGGAGTTCTCTCGCGCACTGTCTACCTGCTGGAGGAGCGGGGTGCCCCAGACCG	840
	75..1952	QY	841	CGGCTGGGGGTTGGGACACTCTCTGTGGGGAGCGGCCAGCCAGCCACTTGTGTGC	900
	/note="unnamed protein product"	Db	841	CGGCTGGGGGTTGGGACACTCTCTGTGGGGAGCGGCCAGCCAGCCACTTGTGTGC	900
	/codon_start=1	QY	901	TACCACGCGCTTCTCTTCCGACTGCTGAGCGGGAGCGGATCGCGACATCGAGCGCC	960
	/protein_id="CAD80012.1"	Db	901	TACCACGCGCTTCTCTTCCGACTGCTGAGCGGGAGCGGATCGCGACATCGAGCGCC	960
	/db_xref="GI:2913661"	QY	961	ACTTGGCTGCAATGTTTCAAGAGCGTGTGAAGCAGGAGGCGCTCGCGTGGGTGCAGGAC	1020
	LSKALLPTALLNVTTRAAPRLQRLSPQCAEVRGSDXDKKYPKFRDRERRA	Db	961	ACTTGGCTGCAATGTTTCAAGAGCGTGTGAAGCAGGAGGCGCTCGCGTGGGTGCAGGAC	1020
	ERAYREKNEITLALCFVFPFIVCTVLRQQLRGLRSLQRAEVLRLSRTSKITTSVYJFITSV	QY	1021	AGGACAGGCGTGGCCGGAGTGCACACAGAGTGCATCGAGGGGGGCAAGGGCTCGAGG	1080
	LSAPVADPRIGQJLNLRLAEGVLRRAQFAKLELQELRGSKVCTJFLSKKE	Db	1021	AGGACAGGCGTGGCCGGAGTGCACACAGAGTGCATCGAGGGGGGCAAGGGCTCGAGG	1080
	LPVLTEVTFDQSFQEFALSYLEDGQVRLTAAGVGTLLRDAQPHSLVJ	QY	1081	ACACGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1140
	TRFLPGLLSAERNRIERHFGVSRVQKALRWQGGCGCPVAPETVGAAGL	Db	1081	ACACGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1140
	EDTEPEEBEENPVPVLELLVCLVETQEDAFVROALCRFPFELALQVRFRCMBVAV	QY	1141	TGTACTGCTGTACGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1200
	LSVCRCCPAGNALRLISRLVAAQEKKSLGKLQASLGGSSQGTTKOLPASLJH	Db	1141	TGTACTGCTGTACGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1200
	PURQANTDPLCHLSLITSHCKLPDRAVCRDLSEALRAAPALTELLHRLSEALRM	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	LSGGJAWPCRVQTVRVQJPDQPGQLYLVGMLRSPATLLDLSGCOLPAPMVTYLC	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	AVLHQCGLQTLSLASVLSBQSLQELQAVKRAKFDLVITHPALDGHPPPKELIST	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	F"	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
BASE COUNT	534 a 861 c 844 g 450 t	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
ORIGIN		Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	Query Match 100.0%; Score 2689; DB 6; Length: 2689;	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	Best Local Similarity 100.0%; Pred. No. 0;	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	Matches 2689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	1 CGACCGTGGCGCGCAGCCTGGCTGACCTGATCTCTCTGACAGTGGCGCGGCGGCG 60	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	1 CGACCGTGGCGCGCAGCCTGGCTGACCTGATCTCTGACAGTGGCGCGGCGGCGG 60	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	61 CGCGGTGGCGCAGAGTCTGCTGCGCCAGCGCGGCTGCTTTCATCTGGAGCGCGG 120	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	61 CGCGGTGGCGCAGAGTCTGCTGCGCCAGCGCGGCTGCTTTCATCTGGAGCGCGG 120	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	121 ACAGCTGCGCGGCTGGGGGGGCGCGAGCGCGCGCTGTCACAGACCGCTTCGAGCGG 180	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	121 ACAGCTGCGCGGCTGGGGGGGCGCGAGCGCGCGCTGTCACAGACCGCTTCGAGCGG 180	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	181 CGAGCGCGCGGGTCTAGCGGGTCTGAGTAAAGGCGTCTGCCACCGGCTCC 240	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	181 CGAGCGCGCGGGTCTAGCGGGTCTGAGTAAAGGCGTCTGCCACCGGCTCC 240	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	241 TGGTGTGACACGCGCGCGCGCGCGCGCGCGCGCTGTGTGTCCCGG 300	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	241 TGGTGTGACACGCGCGCGCGCGCGCGCGCGCGCTGTGTGTCCCGG 300	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	301 AGTGGCGGAGTGGCGGCTTCTCGACAAAGACAAAGAGTATTTCTCAAGTCT 360	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	301 AGTGGCGGAGTGGCGGCTTCTCGACAAAGACAAAGAGTATTTCTCAAGTCT 360	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	361 TCGGAGTGGAGGCGCGGCGCTTACCGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGG 420	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	361 TCGGAGTGGAGGCGCGGCGCTTACCGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGG 420	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	421 TCGGCTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	421 TCGGCTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	481 TGGAGCTGGTGGGAGCTGTGGCGACGCTTCAAGACCAACAGTCAAGTCTACTGCTTT 540	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	481 TGGAGCTGGTGGGAGCTGTGGCGACGCTTCAAGACCAACAGTCAAGTCTACTGCTTT 540	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	541 TCATCACAGGCTTCTAGCTCGGCTCGGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCG 600	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	541 TCATCACAGGCTTCTAGCTCGGCTCGGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCG 600	QY	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260
	601 TGCAGTCTGTGCGGCTGCGCGCGGCGGCGGCTTCTGCGACCGAGCGGCGGCTTGGCG 660	Db	1261	CGGAGCTGGCGCTGACGAGTGGCTTCTGCGCATGGAGTGGCTGTTCTGAGCTACT	1260

Qy	787	AGTTCTCTGGGCACTGTCTCTA	CTGCTGAGAGACGGGCGGGG	TGCCCTAGACACCGCGCTG	846	
Db	1511	AGTTCTCTGGGGCACTGTCTCTA	CTGCTGAGGAGACGGCGGGG	TGCCCTAGACACCGCGCGCTG	1570	
Qy	847	GGGCGTGGGACACTCCTG	TGTGGGACACCCAGCGGCACACGCCA	CTTGGTGTCTCACCA	956	
Db	1572	GGGCGTGGGACACTCCTG	TGTGGGAGCGCCAGCCGACACGCCA	CTTGGTGTCTCACCA	1630	
Qy	907	CGCGCTTCCTCTTGGA	CTGCTGAGCGCGAGCGGAT	TGCGCGACATCGAGCGCACTTCG	966	
Db	1631	CGCGCTTCCTCTTCGGA	CTGAGCGCGGAGCGGAT	TGCGCGACATCGAGCGCACTTCG	1690	
Qy	967	GCTGCATGTTTACAGACG	GTGTGAAGCAGACAGGCCCT	CTGCGGTGGGTGCASGGACAGGGAC	1026	
Db	1691	GCTGCATGTTTCAAGCGT	TGTGTGAAGCAGACAGGCCCT	CTGCGGTGGGTGCAGGGACAGGGAC	1750	
Qy	1027	AGGCTCCCGGAGTGGCA	TGCGATCAGAGGTGACCGAGGGG	GCCTCAAGGGCTCGAGGACACCG	1086	
Db	1751	AGGCTCCCGGAGTGGCA	TGCGATCAGAGGTGACCGAGGGG	GCCTCAAGGGCTCGAGGACACCG	1810	
Qy	1087	AAGAGCCAGAGGAGGAG	AGGAGGAGGAGAGGCCCA	CTACTCCACCTGGAGTTCCTGTACT	1146	
Db	1811	AAGAGCCAGAGGAGGAG	AGGAGGAGAGGCCCA	CTACTCCACCTGGAGTTCCTGTACT	1870	
Qy	1147	GCTGTACGAGACGACG	AGGAGGAGCGCTTTGTGCGCC	AAGGCCCTGTGCGGTTCCCGAGC	1206	
Db	1871	GCTGTACGAGACGACG	AGGAGGAGCGCTTTGTGCGCC	AAGGCCCTGTGCGGTTCCCGAGC	1930	
Qy	1207	TGGCGCTGCAGCGATCG	CTCTGCGCATGAGAGTGGGT	TCTGAGTACTGGGTGA	1266	
Db	1931	TGGCGCTGCAGCGATCG	CTCTGCGCATGAGAGTGGGT	TCTGAGTACTGGGTGA	1590	
Qy	1267	GGTCTGCCCTGCTG	TCGACGCACTCGCGCT	GATCAGCTGCAGATTGGTTGCTGC	1326	
Db	1991	GGTCTGCCCTGCTG	TCGACGCACTCGCGCT	GATCAGCTGCAGATTGGTTGCTGC	2050	
Qy	1327	AGAPAGAGAGAGAGAG	CGCTGGGAGAGCGGCT	CCAGGCGAGGCTGGGTGGCGG	CAGTCTCTC	1386
Db	2051	AGAAGAAGAGAGAGAG	CGCTGGGAGAGCGGCT	CCAGGCGAGGCTGGGTGGCGG	CAGTCTCTC	2110
Qy	1387	AAGSCACCAAAACAA	CTGTCAGCTCCCTCTCT	TCTCATCCACTCTTC	CAGGCAATGACTG	1446
Db	2111	AAGSCACCAAAACAA	CTGTCAGCTCCCTCTCT	TCTCATCCACTCTTC	CAGGCAATGACTG	2170
Qy	1447	ACCCACTGTGCCATCT	GAGCAGGCTCAGGTGT	CTCCACTGCAAACTCCCT	TGACGCGGTCT	1506
Db	2171	ACCCACTGTGCCATCT	GAGCAGGCTCAGGTGT	CTCCACTGCAAACTCCCT	TGACGCGCGGTCT	2230
Qy	1507	GCCGAGACCTTTCT	GAGGCGCTCAGGSC	CAGCCCCCGCACTG	ACGAGAGCTGGGCTCTCTC	1566
Db	2231	GCCGAGACCTTTCT	GAGGCGCTCAGGSC	CAGCCCCCGCACTG	ACGAGAGCTGGGCTCTCTC	2290
Qy	1567	ACACAGGCTCAGTGAG	CGCGGAGCTGGTATG	CTGAGTGAGGCTAGCTG	CGCCGAGT	1626
Db	2291	ACACAGGCTCAGTGAG	CGCGGAGCTGGTATG	CTGAGTGAGGCTAGCTG	CGCCGAGT	2350
Qy	1627	GCAGGCTGCAGACGGT	CAGGTACAGTGTCC	TGACCCCGACGAGG	GCCTCCAGTACCTGG	1686
Db	2351	GCAGGCTGCAGACGGT	CAGGTACAGTGTCC	TGACCCCGACGAGG	GCCTCCAGTACCTGG	2410
Qy	1687	TGGGTATGCTTCGG	CAGAGGCCCCGCCCTG	ACCCCTGGATCTCAG	CGGCTGCGCACTGC	1746
Db	2411	TGGGTATGCTTCGG	CAGAGGCCCCGCCCTG	ACCCCTGGATCTCAG	CGGCTGCGCACTGC	2470
Qy	1747	CCGCCCCCATGGT	CAGCTTCTGTGTG	TCAGTCTCGACAC	CCAGGATGGGGCTG	1806
Db	2471	CCGCCCCCATGGT	CAGCTTCTGTGTG	TCAGTCTCGACAC	CCAGGATGGGGCTG	2530
Qy	1807	CCCTCAGTCTGGCTCT	TGTGAGCTGAGC	GAGTCACTTACAGG	AGCTTCAGCTGTGA	1866
Db	2531	CCCTCAGTCTGGCTCT	TGTGAGCTGAGC	GAGTCACTTACAGG	AGCTTCAGCTGTGA	2590

Qy	1867	AGAGAGCAAAAGCCGATCTGGTCTATCACACACCCAGCGCTGGACGGCCACCCCAACACCTC	1926
Db	2591	AGAGAGCAAAAGCCGATCTGGTCTATCACACACCCAGCGCTGGACGGCCACCCCAACACCTC	2650
Qy	1927	CCAAGGAACCTATCTCGACCTCTGAGGCTCTGGTGGCCAGACAGCGGTGGAGAGACCTTA	1986
Db	2651	CCAAGGAACCTATCTCGACCTCTGAGGCTCTGGTGGCCAGACAGCGGTGGAGAGACCTTA	2710
Qy	1987	GTCAAGTCTCCTGTGGAGAGAAAGGCCCATTTCAAGGGCAGAGAGATATTTGCTCTCGGCC	2046
Db	2711	GTCAAGTCTCCTGTGGAGAGAAAGGCCCATTTCAAGGGCAGAGAGATATTTGCTCTCGGCC	2770
Qy	2047	TTTGGGAAACTTTTGGAGCGAGAGCGGAGACAGGATGTGGGAGGCGCCAGACACGGCA	2106
Db	2771	TTTGGGAAACTTTTGGAGCGAGAGCGGAGACAGGATGTGGGAGGCGCCAGACACGGCA	2830
Qy	2107	CCCTGCCCTCCTCAGGACAGCGCCAGACCTGCCCCCTCTCTCCACACCTCGGGGTACCCCT	2166
Db	2831	CCCTGCCCTCCTCAGGACAGCGCCAGACCTGCCCCCTCTCTCCACACCTCGGGGTACCCCT	2890
Qy	2167	TCTCCCCAGCCGACCTACTTCCACCGACCTTCTCTCTCCTCAGACCTCCAGCCATTC	2226
Db	2891	TCTCCCCAGCCGACCTACTTCCACCGACCTTCTCTCTCCTCAGACCTCCAGCCATTC	2950
Qy	2227	CCCTTGAAACACCCGCCAGCCCAAGCCACAATAATGACAGCGAGAGCTCCCAATTAACT	2286
Db	2951	CCCTTGAAACACCCGCCAGCCCAAGCCACAATAATGACAGCGAGAGCTCCCAATTAACT	3010
Qy	2287	AAGCACTACTTGGCGCAGAAATAACCTTCACTGCTCTGATGCCATCTCAGTGTGGCC	2346
Db	3011	AAGCACTACTTGGCGCAGAAATAACCTTCACTGCTCTGATGCCATCTCAGTGTGGCC	3070
Qy	2347	CAACAGCCCCAGAACTATGCCACATAGACTGGAGTAGGCACTTCAACCTCCCTCCCT	2406
Db	3071	CAACAGCCCCAGAACTATGCCACATAGACTGGAGTAGGCACTTCAACCTCCCTCCCT	3130
Qy	2407	GTTAGGAATGAGACCATCTCTGAGGCTATGGCCAGGCCACAGGGCTCAGTGTCTGAG	2466
Db	3131	GTTAGGAATGAGACCATCTCTGAGGCTATGGCCAGGCCACAGGGCTCAGTGTCTGAG	3190
Qy	2467	ATCTTTGGGAAGGAGACTAGGCGAGTGGAGACAGCGAGACACCCCGTGTGGGTGGG	2526
Db	3191	ATCTTTGGGAAGGAGACTAGGCGAGTGGAGACAGCGAGACACCCCGTGTGGGTGGG	3250
Qy	2527	AAGCATGACCACATGGTGGTGGAGAGCCCCCATGCACTGACGTAATTCCTGTGGA	2586
Db	3251	AAGCATGACCACATGGTGGTGGAGAGCCCCCATGCACTGACGTAATTCCTGTGGA	3310
Qy	2587	CTCATTTCTGGTTCTATTACACTGGCCAGGCGGTGGTACAAATACAGGTGG	2640
Db	3311	CTCATTTCTGGTTCTATTACACTGGCCAGGCGGTGGTACAAATACAGGTGG	3364
RESULT	3		
LOCUS	AX663298	2763 bp	DNA linear PAT 22-MAR-2003
DEFINITION	Sequence 5 from Patent WO02061086.		
ACCESSION	AX663298		
VERSION	AX663298.1	GI:29163662	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Feder, J., Ramanathan, C. and Mintier, G.		
TITLE	Human leucine-rich repeat containing protein, expressed predominantly in small intestine, HLRRS11		
JOURNAL	Patent: WO 02061086-A 5 08-AUG-2002;		
FEATURES	Bristol-Myers Squibb Company (US)		
source	Location/Qualifiers		
	1..2763		
	/organism="Homo sapiens"		

[illegible][illegible]

```
Db 2042 GAGCAGGTTGAGACACCTTAAAGTCCCTGTGGAGAGACGGCCATTCACAGGGCA 2100
Qy 2027 GGAGGATATTGCTCTCGGCCCTTTGGGAACTTTTGAGCCGAGAGGCGCAGACAGGCATG 2086
Db 2101 GGAGGATATTGCTCTCGGCCCTTTGGGAACTTTTGAGCCGAGAGGCGCAGACAGGCATG 2160
Qy 2087 TGGAGGCCCGCAGACACGGCAACCTGCGCCGCTCCAGGACAGGCGCCAGCACTGCTCTC 2146
Db 2161 TGGAGGCCCGCAGACACGGCAACCTGCGCCGCTCCAGGACAGGCGCCAGCACTGCTCTC 2220
Qy 2147 TCCACACCTGGGGTACCCCTTCTCCCCAGCCGCCACCACTACTCCACCCACTTCTCTC 2206
Db 2221 TCCACACCTGGGGTACCCCTTCTCCCCAGCCGCCACCACTACTCCACCCACTTCTCTC 2280
Qy 2207 CTGAGACCCCTCCAGCCATTCCTTGAAGACACCCCGCCAGCCAGCCACCAATATGAC 2266
Db 2281 CTGAGACCCCTCCAGCCATTCCTTGAAGACACCCCGCCAGCCAGCCACCAATATGAC 2340
Qy 2267 AGCAGAGCTCCCAATTAACCTAAGCAGCTTCTGCGGAGAGCTTCTGCGGAGAGCTGA 2326
Db 2341 AGCAGAGCTCCCAATTAACCTAAGCAGCTTCTGCGGAGAGCTTCTGCGGAGAGCTGA 2400
Qy 2327 TCCCATCTGCACTGTGCGCCCAAGCAGCCCGCCAGAACTATGCCACATAGACTGAGGTAG 2386
Db 2401 TCCCATCTGCACTGTGCGCCCAAGCAGCCCGCCAGAACTATGCCACATAGACTGAGGTAG 2460
Qy 2387 GCACTTCACTGCTCCCTCTTGAAGATGAGACCTTCTGCGGAGAGCTTCTGCGGAGAGCTGA 2446
Db 2461 GCACTTCACTGCTCCCTCTTGAAGATGAGACCTTCTGCGGAGAGCTTCTGCGGAGAGCTGA 2520
Qy 2447 ACAGCGTCCAGTCTGAGATCTTTGGGAGGAGAGCTAGGCGAGGTGAGACAGCGCA 2506
Db 2521 ACAGCGTCCAGTCTGAGATCTTTGGGAGGAGAGCTAGGCGAGGTGAGACAGCGCA 2580
Qy 2507 SAACCCCGCTGCTGGGTGGAGAGCATGACCACTATGTTGGGTGAGAGCCCGCTTGCATG 2566
Db 2581 GAACCCCGCTGCTGGGTGGAGAGCATGACCACTATGTTGGGTGAGAGCCCGCTTGCATG 2640
Qy 2567 ACCTTAATTCCTGCTGAGCTATCTTGTGTTTCTATTACACTGCGCGAGGTG 2626
Db 2641 ACCTTAATTCCTGCTGAGCTATCTTGTGTTTCTATTACACTGCGCGAGGTG 2700
Qy 2627 TACAAACAGCTCGGTCTCACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2686
Db 2701 TACAAACAGCTCGGTCTCACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2760
Qy 2687 AAA 2699
Db 2761 AAA 2763
```

```
RESULT 4
AX663299 LOCUS 2054 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 6 from Patent WO02061086.
ACCESSION AX663299
VERSION AX663299.1 GI:29163663
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1
    Feder, J., Ramanathan, C. and Mintier, G.
    Human leucine-rich repeat containing protein, expressed
    predominantly in small intestine, HLRSL1
    JOURNAL Patent: WO 02061086-A 6 08-AUG-2002;
    Bristol-Myers Squibb Company (US)
  FEATURES
    source location/Qualifiers
      1..2054 /organism="Homo sapiens"
      /mol_type="genomic DNA"
```

```
BASE COUNT 473 a 631 c 607 g 343 t
ORIGIN
Query Match 75.1%; Score 2020.2; DB 6; Length 2054;
Best Local Similarity 99.9%; Pred. No. 3.4e-306;
Matches 2022; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 665 GGAACGTGAGCAACTGAGCTTCTGAGCTCCAAAGTGCAGACGCTGTTTCTCAGCAAAA 724
Db 1 GGAACGTGAGCAACTGAGCTTCTGAGCTCCAAAGTGCAGACGCTGTTTCTCAGCAAAA 60
Qy 725 GGAGTCCCGGGCTGCTGGAGACAGAGTCACTACAGTTCATCGACAGAGCTTCCA 784
Db 61 GGAGTCCCGGGCTGCTGGAGACAGAGTCACTACAGTTCATCGACAGAGCTTCCA 120
Qy 785 GGAGTTCCTCGCGGCACTGTCTTACCTGCTGAGAGACGCGGGGTGCCAGGACCGGGC 844
Db 121 GGAGTTCCTCGCGGCACTGTCTTACCTGCTGAGAGACGCGGGGTGCCAGGACCGGGC 180
Qy 845 TGGCGGCTTGGGACACTCTGCTGGGAGCCCGACGCGGACAGCTTGTGCTCAC 904
Db 181 TGGCGGCTTGGGACACTCTGCTGGGAGCCCGACGCGGACAGCTTGTGCTCAC 240
Qy 905 CACGCGCTTCTCTTGGAGCTGCTGAGCGAGCGGATGCGGACATCGAGCGCCACTT 964
Db 241 CACGCGCTTCTCTTGGAGCTGCTGAGCGAGCGGATGCGGACATCGAGCGCCACTT 300
Qy 965 CGGCTGCATGTTTCAAGCGCTGTGAAGCAGAGCGCTCGGTTGGGTGACAGGACAGGG 1024
Db 301 CGGCTGCATGTTTCAAGCGCTGTGAAGCAGAGCGCTCGGTTGGGTGACAGGACAGGG 360
Qy 1025 ACAGGCTGCCCGGAGTGGCCAGAGGTGACCGAGGGGGCCAAAGGCTTCGAGGACAC 1084
Db 361 ACAGGCTGCCCGGAGTGGCCAGAGGTGACCGAGGGGGCCAAAGGCTTCGAGGACAC 420
Qy 1085 CGAAGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1144
Db 421 CGAAGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Qy 1145 CTGCTGTACAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1204
Db 481 CTGCTGTACAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Qy 1205 GCTGCGGCTGAGCGAGTGGCTTCTGCGGATGAGAGTGGCTTCTGAGCTACTCGGT 1264
Db 541 GCTGCGGCTGAGCGAGTGGCTTCTGCGGATGAGAGTGGCTTCTGAGCTACTCGGT 600
Qy 1265 GAGTGTGCTGCTGTGACAGGCACTGCGGCTGATCAGCTGACAGATTGGTTGCTGGCA 1324
Db 601 GAGTGTGCTGCTGTGACAGGCACTGCGGCTGATCAGCTGACAGATTGGTTGCTGGCA 660
Qy 1325 GGAGAAGAAGAGAGCTTGGGGAAGCGGCTCCAGGCGAGCGCTGGGTGGCGGAGTTTC 1384
Db 661 GGAGAAGAAGAGAGCTTGGGGAAGCGGCTCCAGGCGAGCGCTGGGTGGCGGAGTTTC 720
Qy 1385 TCAAGGACCAACAAACAACTCCAGCTTCTTCTATCCACTCTTTTTCAGGCAATGAC 1444
Db 721 TCAAGGACCAACAAACAACTCCAGCTTCTTCTATCCACTCTTTTTCAGGCAATGAC 780
Qy 1445 TGACCCACTGTGCCATCTGAGCAGCTCACGCTGTGCCACTGCAAACTCCCTGACGGGT 1504
Db 781 TGACCCACTGTGCCATCTGAGCAGCTCACGCTGTGCCACTGCAAACTCCCTGACGGGT 840
Qy 1505 CTGCGGAGACCTTTCTGAGGCGCTGAGGGGAGCGCGGCTGACGAGCTGGGCGCTCT 1564
Db 841 CTGCGGAGACCTTTCTGAGGCGCTGAGGGGAGCGCGGCTGACGAGCTGGGCGCTCT 900
Qy 1565 CCACAAACAGGCTCAGTGGAGGCGGAGCTGCGTATCTGAGTGAGGCGCTAGCTCTGCCGCA 1624
Db 901 CCACAAACAGGCTCAGTGGAGGCGGAGCTGCGTATCTGAGTGAGGCGCTAGCTCTGCCGCA 960
Qy 1625 GTGAGGGGTGACAGCGGTCAAGGTACAGCTGCTGACGCCCCCAGCGGGGCTTCCAGTACCT 1684
```

```
Db 961 GTCCAGGAGCGACACGGTCAGGAGTACAGTGGCTGACCCCGAGGAGGCTCCAGTACCT 1020
Qy 1685 GGTGGGTATGCTTCGGCAGAGCCCCCGCTGACCCAGCTGATCTTCAGCGGCTGCGCAACT 1744
Db 1021 GGTGGGTATGCTTCGGCAGAGCCCCCGCTGACCCAGCTGATCTTCAGCGGCTGCGCAACT 1080
Qy 1745 GCCCGCCCGATGGTGAACCTAAGTGTGGAGTCTTCAGCAGCAGGAGTGGCGCTGCA 1804
Db 1081 GCCCGCCCGATGGTGAACCTAAGTGTGGAGTCTTCAGCAGCAGGAGTGGCGCTGCA 1140
Qy 1805 GACCTCAGTCTGGCTCTGTGGAGTCTGGAGGAGTCTTCAGCAGGAGTCTTCAGGCTGT 1864
Db 1141 GACCTCAGTCTGGCTCTGTGGAGTCTGGAGGAGTCTTCAGCAGGAGTCTTCAGGCTGT 1200
Qy 1965 GAAGAGAGCAAGCCGATCTGTGTATCAGCAGCAGGCTGAGCGGCGCCACCAACC 1924
Db 1201 GAAGAGAGCAAGCCGATCTGTGTATCAGCAGCAGGCTGAGCGGCGCCACCAACC 1260
Qy 1925 TCCAGAGGACTCATCTGACCTTCTGAGGCTCTGGTGGCAGAGCAGGCTGGAGAGCC 1984
Db 1261 TCCAGAGGACTCATCTGACCTTCTGAGGCTCTGGTGGCAGAGCAGGCTGGAGAGCC 1320
Qy 1985 TAGTCAAGTCCCTGTGGAGAGAACCGCCCATTCACAGGCGCAGGAGATATGCTCTGG 2044
Db 1321 TAGTCAAGTCCCTGTGGAGAGAACCGCCCATTCACAGGCGCAGGAGATATGCTCTGG 1380
Qy 2045 CTTTGGGAACTTTTGGAGCCAGAGCGCGCAGAGCAGGATGTGGAGGCGCCAGACAGG 2104
Db 1381 CTTTGGGAACTTTTGGAGCCAGAGCGCGCAGAGCAGGATGTGGAGGCGCCAGACAGG 1440
Qy 2105 CACCTGCGCCGTCAGGAGACA3CCCGAGAGCCTGCCCTCTCCACACCTGGGGTACCC 2164
Db 1441 CACCTGCGCCGTCAGGAGACA3CCCGAGAGCCTGCCCTCTCCACACCTGGGGTACCC 1500
Qy 2165 CTTCTCCCCAGCCCACTACTCCACCCAGCTTCTCTCTCTGAGACCTTCCAGCCAT 2224
Db 1501 CTTCTCCCCAGCCCACTACTCCACCCAGCTTCTCTCTCTGAGACCTTCCAGCCAT 1560
Qy 2225 TCCCTTGGAAACCCCGCCGAGCCCAAGCCCAAGTAAATGAGAGGAGAGTCCCAATAA 2284
Db 1561 TCCCTTGGAAACCCCGCCGAGCCCAAGCCCAAGTAAATGAGAGGAGAGTCCCAATAA 1620
Qy 2285 CTAGACACCTACCTGCGGCGAGAAATACCTTCACTCCCTGATCCCATCTGCAGTGTGG 2344
Db 1621 CTAGACACCTACCTGCGGCGAGAAATACCTTCACTCCCTGATCCCATCTGCAGTGTGG 1680
Qy 2345 CCAACAGCCCGCAGAACTATGCCACATAGACTGAGGATAGGAGTCCAGCTCCCTCC 2404
Db 1681 CCAACAGCCCGCAGAACTATGCCACATAGACTGAGGATAGGAGTCCAGCTCCCTCC 1740
Qy 2405 CTGTTAGGAATGAGACATCCCTGAGGCTATGGCCAGGCGCCAGAGGCTCCAGTGTCTG 2464
Db 1741 CTGTTAGGAATGAGACATCCCTGAGGCTATGGCCAGGCGCCAGAGGCTCCAGTGTCTG 1800
Qy 2465 AGATCTTTGGAGAGGAGACTAGGAGGAGTGGAGAGCAGCGCAACCCCGTGTCTGGTG 2524
Db 1801 AGATCTTTGGAGAGGAGACTAGGAGGAGTGGAGAGCAGCGCAACCCCGTGTCTGGTG 1860
Qy 2525 GGAAGCATGACCATGAGTGGTGGTGGAGCGCCCGCTGCACTGACGCTAAATCCCGCTGTG 2584
Db 1861 GGAAGCATGACCATGAGTGGTGGTGGAGCGCCCGCTGCACTGACGCTAAATCCCGCTGTG 1920
Qy 2585 GACTCATTTCTGTTGTTCTTATTAACCTTGGCAGGCGTGGTACAAATACAGGTCGGTGC 2644
Db 1921 GACTCATTTCTGTTGTTCTTATTAACCTTGGCAGGCGTGGTACAAATACAGGTCGGTGC 1980
Qy 2645 TCACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2689
Db 1981 TCACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2025
```

```
AF479748 2679 bp mRNA linear PRI 12-AUG-2002
LOCUS AF479748 Homo sapiens PYRIN-containing APAF1-like protein 5 mRNA, complete cds.
DEFINITION AF479748 Homo sapiens PYRIN-containing APAF1-like protein 5 mRNA, complete cds.
ACCESSION AF479748
VERSION AF479748.1
KEYWORDS GI:19387135
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE Wang,L., Manji,G.A., Grenier,J.M., Al-Garawi,A., Merriam,S., Lora,J.M., Geddes,B.J., Briskin,M., DiStefano,P.S. and Bertin,J.
TITLE PYPAF7, a Novel PYRIN-containing APAF1-like Protein: That Regulates Activation of NF-kappa B and Caspase-1-dependent Cytokine Processing
J. Biol. Chem. 277 (33), 29874-29880 (2002)
PUBMED 12019269
REFERENCE 2 (bases 1 to 2679)
AUTHORS Bertin,J.
TITLES Direct Submission
JOURNAL Submitted (01-FEB-2002) Inflammation, Millennium Pharmaceuticals Inc., 45 Sidney Street, Cambridge, MA 02139, USA
FEATURES
Location/Qualifiers
1..2679
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosomes="11"
/maps="11p15"
1..2679
/note="PYPAF5; PYRIN-containing nucleotide-binding site/leucine-rich repeat family member"
/product="PYRIN-containing APAF1-like protein 5"
/protein_id="AAL87105.1"
/db_xref="GI:19387136"
/translation="MDQPEAFCSSTGPRLAVARELLAALBELSDELKSRHKLKRDV
GDGSRIPWGLERADAVDLAEQAFYGPPEALEVARKTLKRADRDVAQLCERRL
QRLGSGLLSVSEYKKYREHLQRLHARVKNARSVKTKRKLGLIAPESAAPE
EALGABEPEPGARRSDTHFNRLFRDEGRRLTVLQGFAGITKTAAKILYD
WAAGLYQOVDFAFMPGCLLERLSADLIDQCPDRGAPVPMQACQPLLF
ILDGDELPAFGPEAAPCTDPEAASGARVLGLLSKALLPTALLVTTAAAPGLR
QRLCSPOCAEVRGSDKDKYFYKFRDERAERAVRVKENSTLFCALCFVPVCW
IVCTVLRQLELGRDLRSKTSKTSVYLLFITSLSAPVADGPRLOGDLNKLCLAR
EGVLRRAQFAKELEQLERGSKVQTLFLSKESPGVLETEVTYQFIDQSFQELAA
LSYLLDGGVPRTAAGGVTLLRQDAPHSLSVLTFRFLGLLSAERMRDIERHFGCY
VSEVRQKALRVQGGCGGCGVAPETVGAKGLEDTEPEEEEGESEPYPLELLYC
LYETQEDAFVRQALCFPELALQVRFCMDVAVLSYCVRCPPAGQURLL-SCRIVAA
QEKKKSLKRLQASLGSSSGSTGKTQKLPASLLHPLFQAMTDPLCHLSLTLSHCKL
PDACRDLSEALRAAPALTLEGLHNLSEAGRLMLSEGLAWPCQRTVQVQ.PDPQ
RGLQYLVGMRLQSPALFTLDLSGCOLPAPMTYLCVILQHQCGLQTLASVLSBELQ
SLOELQAVKRAKPDLVITHPALDGHQPQKELISTF"
BASE COUNT 458 a 878 c 934 g 409 t
ORIGIN
Query March 71.6%; Score 1925; DB 9; Length 2679;
Best Local Similarity 99.6%; Pred. No. 2.5e-291;
Matches 1941; Conservative 0; Mismatches 5; Indels 3; Gaps 1;
Qy 7 CGTGGGCGGCGAGCTGGCTGACCTGATCCTGGACAGTGGCCCCCGAGCGCGCGCGG 66
Db 731 CGGGCAGCGCGAGCTGGCTGACCTGATCCTGGACAGTGGCCCCCGAGCGCGCGG 790
Qy 67 TSCCGCAGATGCTGGCCCGAGCGCGCTCTCTTCTCATCTGGACCGCGGAGCGAGC 126
Db 791 TGCCTGAGATGCTGGCCCGAGCGCGCTCTCTTCTATCTGGACCGCGGAGCGAGC 850
Qy 127 TGCCTGCGCTGGGGGCGCGCGCGCTCTGACAGACCCCTTCGAGGCGCGGAGCG 186
Db 851 TGCCTGCGCTGGGGGCGCGCGCGCTCTGACAGACCCCTTCGAGGCGCGGAGCG 910
```


gene	/mol_type="mrna" /db_xref="taxon:9606" 1..2679 /gene="NALP6"	
CDS	1..2679 /gene="NALP6" /note="member of the NALP protein family involved in inflammation; contains NACHT, LRR and PYD domains" /codon_start=1 /product="NALP6" /protein_id="AAO18157.1" /db_xref="GI:28436366" /translation="MDQPEARCSSTGPRFLAVARELLLAALIELSQBELKFRKFLRDV GPDGSLGWLKERADADVLAELQAQFYGPPEPALEVARKTLKRADARDVAQQLGERRL QRLGSLGTLTVSVSYKKYREHVLQDHARVYKERNASVKITRFTSLJLJAPESAPE EALGPEPEGRARSDTHFNRLFRDEGRRLPVLQGGPAGIGKTKAAKSIYD WAAGKLYQGVJFAFPMFGSLLERPGTRSLADILQCDRGPAPVPMIAQPCRLLF ILGDADELPAUGPEAAPCTDPFAAGSARVLGGLSKALLPTALLLVTTAAAPQRL QRLCSPOCAEVRGSDKDKKFKYKFFERRERAERAVRFVKENETLFAICFYVPVCM IUCTVLRLQELGRDLSTSKTTSVLLFFTSVLSSNPVADPRLOGDLRLNCLRAK EVLGRRQAPEKLELQELGSKVQTLFLSKKELPGLVETVYQIDQSDFEFLAA LSYLJEDGVPRPAAGVGTLLRDAQPHSLVLTFRFLGLSAEMRDIERHFOCM VSEYKQFALRWQOQGGCPVAPFVTEGAKLEDTEPEEEBEGEPNYPELLYAC LYETQEDAFKVRQFELALQVRFRVMDVAVLSVCRCPAGQALRLISRLVAA QEKXKSLKLOASLGSSGGTTLKOLPASLLHPLFOAMDPTPLCHLSLTLSHCKL PDACRDLSEALRAAPALTEGLLHNLSEAGLPMISEGLAWPCRVTVYVCLPDPQ RGLQVLVGMKLSQSNALTLIDLSGQQLPAPMWTYLCVLQHQGGLQTLISLASELSEQ SLQELQAVRKARPLDVLITPDLGHPQPKELISTP"	
BASE COUNT	458 a	878 c 534 g 409 t
OR:GIN	Query Match: 71.6%; Score 1925; DB 9; Length 2679; Best Local Similarity 99.6%; Fred. No. 2.5e-291; Matches 194i; Conservative 0; Mismatches 5; Indels 3; Gaps 1;	
QY	7	CQTGGGCGGAGCTGGTGACTGATCTTGGACAGTGCCCGACCGCGGCGCGCGG 66
DB	731	CGGCGACGCGAGCTGGCTGACCTGATCTGGACACAGTGCCCGACCGCGCGCGG 790
QY	67	TGCGCGAGATGTCGCCAGCGCGAGCGGCTCTCTTATCTGACGCGCGGACGAGC 126
DB	791	TGCGCGAGATGTCGCCAGCGCGAGCGGCTCTCTTATCTGACGCGCGGACGAGC 553
QY	127	TGCGCGCGCTGGGGGCGCGGAGCGCGGCTCTGACAGACCCCTTCGAGCGCGAGCG 186
DB	851	TGCGCGCGCTGGGGGCGCGGAGCGCGGCTCTGACAGACCCCTTCGAGCGCGAGCG 910
QY	187	GCGCGGGTGCTAGCGGGTGCTGAGTAGGGGCTGCTGCTGCCAGCGCGCTGCTGG 246
DB	911	GCGCGGGTGCTAGCGGGTGCTGAGTAGGGGCTGCTGCTGCCAGCGCGCTGCTGG 973
QY	247	TGACACGCGCGCGCGCGCGCGCGCGCTGAGGAGGCTGACGAGGCGCGCTGTGT 326
DB	971	TGACACGCGCGCGCGCGCGCGCGCGCTGAGGAGGCTGACGAGGCGCGCTGTGT 1030
QY	307	CCGAGGTGCGCGGCTCTCCGCAAGGACAGAGAGATTTCTACAGTTCTTCGGG 366
DB	1031	CCGAGGTGCGCGGCTCTCCGCAAGGACAGAGAGATTTCTACAGTTCTTCGGG 1093
QY	367	ATGAGAGAGGCGCGAGCGCGCTACCGGCTTCGTGAAGAGAGAGAGAGCTGTTCG 426
DB	1091	ATGAGAGAGGCGCGAGCGCGCTACCGGCTTCGTGAAGAGAGAGAGAGCTGTTCG 1150
QY	427	TGTGCTTCGCGCTTCGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 486
DB	1151	TGTGCTTCGCGCTTCGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1210
QY	487	TGGTGGGACCTGTGCGCGAGCTCCAGAGACCAACAGCTGAGTACCTGCTTTTCAT 546
DB	1211	TGGTGGGACCTGTGCGCGAGCTCCAGAGACCAACAGCTGAGTACCTGCTTTTCAT 1270
QY	547	CCAGGCTTCAGGCTGGCTCCGGTACGCGAGCGGCGCGCGCTTGCAGGCGACCTGG 606

QY	1594	TGGTGGGTATGCTTCGGCAGAGCCCCCGCCTTGACCAACCTTGGATCTCAGCGGCTGCCAAC	1743
DB	2411	TGTTGGGTATGCTTCGGCAGAGCCCTGCCTTGACCAACCTTGGATCTCAGCGGCTGCCAAC	2470
QY	1744	TGCCCGGCCCCCATGGTGACCTTACCTGTGTGTGAGTCTTCGACGACACGAGGATTCGGGCTTGC	1803
DB	2471	TGCCCGGCCCCCATGGTGACCTTACCTGTGTGTGAGTCTTCGACGACACGAGGATTCGGGCTTGC	2530
QY	1804	AGACCCCTCAGTCTGGCCTCTCTGGAGCTGAGCGAGCAGTCACTACAGAGCTTCAGAGCTG	1863
DB	2531	AGACCCCTCAGTCTGGCCTCTCTGGAGCTGAGCGAGCAGTCACTACAGAGCTTCAGAGCTG	2590
QY	1864	TGAAGAGAGCAAAAGCCGATCTGGTCACTACACACCCAGCGCTGGACGGCCACCCACAAC	1923
DB	2591	TGAAGAGAGCAAAAGCCGATCTGGTCACTACACACCCAGCGCTGGACGGCCACCCACAAC	2650
QY	1924	CTCCCAAGGAACCTCATCTCGACCTTCTGA	1952
DB	2651	CTCCCAAGGAACCTCATCTCGACCTTCTGA	2679
RESULT 7			
AX088215			
LOCUS	AX088215	5553 bp	DNA linear PAT 17-MAR-2001
DEFINITION	Sequence 10 from Patent WO114564.		
ACCESSION	AX088215		
VERSION	AX088215.1	GI:13397126	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Peyman, J.A., Green, C.D., Hsu, A., Browning, J.A. and Cardelli, J.		
TIME	Polynucleotides expressed in activated T-lymphocytes and proteins encoded thereby		
JOURNAL	Patent: WO 0114564-A 10 01-MAR-2001;		
FEATURES	Curagen Corporation (US); BIGEN, INC. (US); Location/Qualifiers		
SOURCE	1..5553		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	912 a 1856 c 1879 g 906 t		
ORIGIN			
Query Match	58.4%;	Score 1571.2;	DB 6; Length 5553;
Best Local Similarity	89.6%;	Pred. No. 4.6e-236;	
Matches 1772;	Conservative 0;	Mismatches 38;	Indels 168; Gaps 1;
QY	7	CGTGGGCGCGCAGCCTGGCTACCTGATCTTGGACCAAGTCCCGCAGCGCGGCGCGCGCGG	66
DB	731	CGGCGCAGCGCAGCCTGGCTACCTGATCTTGGACCAAGTCCCGCAGCGCGGCGCGCGCGG	730
QY	67	TGCCCGCAGATGCTGGCCCGAGTCGCGAGCGGGTGTCTTTCATCTTGSACGGCGCGGACGAGC	126
DB	791	TGCCCGCAGATGCTGGCCCGAGTCGCGAGCGGGTGTCTTTCATCTTGSACGGCGCGGACGAGC	830
QY	127	TGCCCGGCGCTGGGGGGGCGCCGAGCGCGGCGCTTGCACAGACCCCTTCGAGCGCGCGAGCG	186
DB	851	TGCCCGGCGCTGGGGGGGCGCCGAGCGCGGCGCTTGCACAGACCCCTTCGAGCGCGCGAGCG	910
QY	187	GCGCGCGGGTGTAGGGGGGTGTGTAGTAGGGCGTGTCTGCCACCGGCGCTCTCTCTGTG	246
DB	911	GCGCGCGGGTGTAGGGGGGTGTGTAGTAGGGCGTGTCTGCCACCGGCGCTCTCTCTGTG	970
QY	247	TGACCAAGCGCGCGCGCGCGCGCGCGGAGGCTTGCAAGGCGCGCTGTGTTCCCGCAGTGC	306
DB	971	TGACCAAGCGCGCGCGCGCGCGCGCGGAGGCTTGCAAGGCGCGCTGTGTTCCCGCAGTGC	1030
QY	307	CCGAGGTGCGCGGCTTCTCCACAAAGGACAAAGAAGTATTCTACAAGTTCCTCCGG	366

Db	1031	CCGAGGTGCGCGCTTCTCCGACAAGGACAAGAAAGTATTTCTACAAGTTCTTCTCGGG	1090
QY	367	ATGAGAGAGGCGCGAGCGCGCTTACCGCTTCGTGAGGAGAGACGACAGCCTGTTCGCGC	426
Db	1091	ATGAGAGAGGCGCGAGCGCGCTTACCGCTTCGTGAGGAGAGACGACAGCCTGTTCGCGC	1150
QY	427	TGTGCTTCTGTGCCCTTCTGTGTCTGGAATCGTGTGCACCGTCTGCGCCAGCAGCTGGAGC	486
Db	1151	TGTGCTTCTGTGCCCTTCTGTGTCTGGAATCGTGTGCACCGTCTGCGCCAGCAGCTGGAGC	1210
QY	487	TCGCTCGGACCTGTCCGGCAGCTCCAAGACACCCAGTCAGTGTACTCGTTTTCATCA	546
Db	1211	TCGCTCGGACCTGTCCGGCAGCTCCAAGACACCCAGTCAGTGTACTCGTTTTCATCA	1270
QY	547	CCAGCGTCTCAGACTCGGCTCCGGTACGCCAGCGGCCCGGTTGCAGGGCGACTGCGCA	606
Db	1271	CCAGCGTCTCAGACTCGGCTCCGGTACGCCAGCGGCCCGGTTGCAGGGCGACTGCGCA	1330
QY	607	ATCTGTCCGCCCTGGCCCGGAGGGGCTCTCGGACGACAGGCGCAGTTTTCAGGCAAAAGG	666
Db	1331	ATCTGTCCGCCCTGGCCCGGAGGGGCTCTCGGACGACAGGCGCAGTTTTCAGGCAAAAGG	1390
QY	667	AATCGAGACAATCGAGCTTCTGGCTCCAAAGTGCAGACGCTGTTCTCAGGCAAAAGG	726
Db	1391	AATCGAGACAATCGAGCTTCTGGCTCCAAAGTGCAGACGCTGTTCTCAGGCAAAAGG	1450
QY	727	AGCTGCCGGCGTCTCGACACAGAGTCACTACCACTTCATCGACACAGACTTCCAGG	786
Db	1451	AGCTGCCGGCGTCTCGACACAGAGTCACTACCACTTCATCGACACAGACTTCCAGG	1510
QY	787	AGTTCCTCGCGCACTGTCTCTACTCTGTGAGGACGCGCGGGTGCACGACCCGCGCTG	846
Db	1511	AGTTCCTCGCGCACTGTCTCTACTCTGTGTGAGGACGCGCGGGTGCACGACCCGCGCTG	1570
QY	847	CGCGGTTGGGACACTCTGCGTGGGACGCCACGCGCACAGCCACTTGGTGTCAACCA	906
Db	1571	CGCGGTTGGGACACTCTGCGTGGGACGCCACGCGCACAGCCACTTGGTGTCAACCA	1630
QY	907	CGCGCTTCTCTTCGAGCTCTGAGCGCGAGCGGATGCGCGACATCGAGCGCACCTTCG	966
Db	1631	CGCGCTTCTCTTCGAGCTCTGAGCGCGAGCGGATGCGCGACATCGAGCGCACCTTCG	1690
QY	967	GCTGCATGTTTTCAGAGCGTGTGAAGCAGAGAGCCCTCGCGTGGTGCAGGGACAGGGAC	1026
Db	1691	GCTGCATGTTTTCAGAGCGTGTGAAGCAGAGAGCCCTCGCGTGGTGCAGGGACAGGGAC	1750
QY	1027	AGGCTGCCCGGAGTGGACACAGAGTGAACGAGGGGGCCAAAGGCTCGAGGACACCG	1086
Db	1751	AGGCTGCCCGGAGTGGACACAGAGTGAACGAGGGGGCCAAAGGCTCGAGGACACCG	1810
QY	1087	AAGAGCCAGAGGAGGAGGAGGAGGAGGAGCCCACTACCCACTGGAGTTCTGTACT	1146
Db	1811	AAGAGCCAGAGGAGGAGGAGGAGGAGGAGCCCACTACCCACTGGAGTTCTGTACT	1870
QY	1147	GCCTGTACGAGCGCAGGAGCGGTTTGTGCGCCAAAGCCCTGTGCGGTTCCCGAGC	1206
Db	1871	GCCTGTACGAGCGCAGGAGCGGTTTGTGCGCCAAAGCCCTGTGCGGTTCCCGAGC	1930
QY	1207	TGGCGCTGACAGTGGCTTCTGCGCATGAGACGTGGCTGTCTGTAGCTACTGCGTGA	1266
Db	1931	TGGCGCTGACAGTGGCTTCTGCGCATGAGACGTGGCTGTCTGTAGCTACTGCGTGA	1990
QY	1267	GGTGTGCCCTGCTGGACAGGCACTGGGCTGATCAGCTGCAGATTGGTTGCTGCGCAGG	1326
Db	1991	GGTGTGCCCTGCTGGACAGGCACTGGGCTGATCAGCTGCAGATTGGTTGCTGCGCAGG	2050
QY	1327	AGAAGAGAAAGAGAGCTGGGGAAGCGGCTCCAGCGCCAGCCTGGTGGCGGAGTTCTC	1386
Db	2051	AGAAGAGAAAGAGAGCTGGGGAAGCGGCTCCAGCGCCAGCCTGGTGGCGGAGTTCTC	2110
QY	1387	AAGGCACCAAAACAACTGCCAGCCTCCCTTCTTCATCCACTCTTT-----	1433
Db	2111	TGGGAGCCCAACTGGGTCCAGAGTAGTACCTTTTCGACACCCCTGCTGTGACATCTGCCCA	2170

[illegible]

```
Db 1871 GCCTGTACGAGACGAGGAGGAGCGTTTGTGGCCAGACCGCTGTGCGGTTCGCGAGC 1930
Qy 1207 TGGCGCTGCAGGAGTGGCTTCTGCCCGATGGAGTGGCTGTCTGAGCTACTGGCTGA 1266
Db 1931 TGGCGCTGCAGGAGTGGCTTCTGCCCGATGGAGTGGCTGTCTGAGCTACTGGCTGA 1990
Qy 1267 GGTGTGCTGCTGTGCAGGAGTGGCTGATCAGCTGCAGATGGTGTCTGCGGAGG 1326
Db 1991 GGTGTGCTGCTGCAGGAGTGGCTGATCAGCTGCAGATGGTGTCTGCGGAGG 2050
Qy 1327 AGAAGAAGAAAGAGCTGGGGAAGCGGCTCCAGGCCAGCGCTGGTGGCGGAGTCTC 1386
Db 2051 AGAAGAAGAAAGAGCTGGGGAAGCGGCTCCAGGCCAGCGCTGGTGGCGGAGTCTC 2110
Qy 1387 AAGGACACACAAACATGCCAGCTCCCTCTCTCATCAGCTCTT 1433
Db 2111 TGGGAGCCCAATGCTGCTCCAGAGTACCTTTCGACCAACCCCTGTGTGACATCTGCCCA 2170
Qy 1434 ----- 1433
Db 2171 CACCTCCACAGACCTCGCTCCTCAGGGCAAGGCTTTCGACAGATTCCTTTGAATA 2230
Qy 1434 ----- 1433
Db 2231 TAGCTCCAAATCAGCCCTCGCCAGGGCTTGGCATCTGTTGAGAGATGAATGTCAAG 2290
Qy 1434 ----- CAGGCAATCACTGACCCAGTGTCC 1458
Db 2291 TGTGGCAGGGCTGGGCTGGGACCCAAAGACCCATGCAATGACTGACCCACTSGCC 2350
Qy 1459 ATCTGAGCAGCTCAGCTGTCCCACTGCACAACTCCCTGAGCGGCTGTCCGAGACCTTT 1518
Db 2351 ATCTGAGCAGCTCAGCTGTCCCACTGCACAACTCCCTGAGCGGCTGTCCGAGACCTTT 2410
Qy 1519 CTGAGCCCTGAGGCGAGCCCGGCTGAGCGAGCTGGGCTCTCCACAAAGGCTCA 1578
Db 2411 CTGAGCCCTGAGGCGAGCCCGGCTGAGCGAGCTGGGCTCTCCACAAAGGCTCA 2470
Qy 1579 GTGAGGCGGACTCGTATGCTGAGTGAAGGCTGAGCTGAGCGAGCTGGGCTCTCCAG 1638
Db 2471 GTGAGGCGGACTCGTATGCTGAGTGAAGGCTGAGCTGAGCGAGCTGGGCTCTCCAG 2530
Qy 1639 CGGTGAGGCTACAGCTGCTGACCCCGGCTGAGCGAGGCTCCAGTACCTGGTGGTATGCTTC 1698
Db 2531 CGGTGAGGCTACAGCTGCTGACCCCGGCTGAGCGAGGCTCCAGTACCTGGTGGTATGCTTC 2590
Qy 1699 GGCAGAGCCCGCTGACCACTGATCTGAGCTGAGCGGCTGCACTGCGCCCGCCCATGG 1758
Db 2591 GGCAGAGCCCTGCGCTGACCACTGATCTGAGCTGAGCGGCTGAGCTGCGCCCGGCTGAG 2850
Qy 1759 TGACCTACCTGTGTGCTGCTGAGGACAGGAGTGGGCTGAGCGGCTGAGCTGAGCT 1816
Db 2851 TGACCTACCTGTGTGCTGCTGAGGACAGGAGTGGGCTGAGCGGCTGAGCTGAGCT 2708
```

```
RESULT 9
LOCUS AX088208
DEFINITION Sequence 3 from Patent WO0114564.
ACCESSION AX088208
VERSION AX088208.1
KEYWORDS Gi:13397119
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiridae; Homo.
1
REFERENCE
AUTHORS Peyton,J.A., Green,C.D., Hsu A., Browning,J.A. and Carulli,J.
TITLE Polynucleotides expressed in activated T-lymphocytes and proteins
encoded thereby
JOURNAL Patent: WO 0114564-A 3 01-MAR-2001;
FEATURES Curagen Corporation (US); BIOGEN, INC. (US)
Location/Qualifiers
```

source

CDS

```
1..6461
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..5556
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC34689.1"
/db_xref="GI:13397120"
/translation="MDQEPACSGTPRLAVARELLAALEBELSQELKRFHRLRDV
GPDGSIWGRLEADADVLAEQAFYGPPEALVARKTKLADARDVAALQERRJ
EALGSGEPLLSVSEYKXKREHVLQHLARVKNARSVKITKRTKLIIAPESRAPE
EALGPAEPEPGARSDTHFNRLFRDEEGRPLTVVLOGPAGIGKMAAKQILYD
WAAGKLYQGVDFAFMPCEGELERPTKSLADLDDQCPDRGAFVPMQAQFQRLIF
ILDGADELPAAGPEAAPTDFEAGSARVGLLSKALLPTALLLLVTTTAAAFGRLL
QGRLSQPAQAEVRFSDKKKRYKFFDERRAERAYFVKNETLFAICFVFFVCW
1VCTVLRQQLGLRDLSTKTTTSVYLLFITSVLSSAPVADGPRQLQGLNCLRLAR
EGVLRAGQAFKELEQLRGSKVOTLFLSKKELPGVLETVYQFIDQSCFELAA
LSVLLDGGVPRTAAGGCGTLLRGAQPHSHVLTRELGLLSAERWRDIERHFGCM
VSRVYKQEAALRWQGGCGGPGVAFVETEGAKGLEDEPEPEEEREEENYFLELLYC
LLETDBAFVQALCRFPFELALQRFRCMDVAVLSYCVRCPPAGQAUURLISRLVAA
QEKXKSLGRLQASLGGSGMLGTQLAPEVPPPCDDICTPTPPDLLQKGAARV
PLNIAP10PLPRGLASVERMNVTLAGAPGDPKTHAMTDPLCHLSSITLSCKLPDA
VCDLSEALRAAPALTLEGLLNRLSEAGLRLMSEGLAWPQCRVOTVRVLPDPORGL
QYLVGMRLQSPALTTDLSSGCOLPAPMTYLCVLOHOGCGLOTLISLSPDPTSSFF
SCRCPEPRRLGLESRMPRSPAPESGDSARTQVEAAGAGGGRRRRPPRPPGHPQ
PRDGAARGPGSFHSGRFVQGTGPRTPKTPRTPAGTEGSGRGRGRESTSRPRPSPD
RPRRTGPASORPPGSGRPRTFLVAPQGSFSLPALAWSRGTVPTLAPGRVG
LRLPLRSPMEDAGEDPTTAAHSLPSDPRLIATVNTAYLTVRFDHLHVSQVYAH
GGDTHRALPSPNVRLEAPAGMGQLTETALDNTGSLFHLTSGPRFRASQCIYAH
RTLPHVLAFRVSIARLAPSGPTILLRSPSPESPDLDLHOGPDPFGARVLYGTLT
PSPQGPQOEVHM:WTPAPDPLTGLGEARTMDFLTAVGSSQAEQAQALTEALQQA
RQALYTAQAQAWAQLWECGLDVLPQLRQALRSLYLLSALPQAPAGTICGSL
PGLSNGREBCTYGHVFDWDLMPFPVSLMHPPEAARAL-EYRITRTDGLALEAONL
GYQKAFWESADSLVCPEDYGVQEVHNVGAVVLAPELYHTTDLQLFREAGN
DIVRVAIFWCSREVSPEEKYHVLGMSPEYHSGVNNVYTNVLVONSRLFAAL
ADOLGPTFSOW:LAVDKIKVPDFEONFHPDFDGVDPDPCVCPSSORLHPGVEV
VQAQADVLLGYVPFSLSPDVRKNLEIYEAVTSQQGPAMTWSK:FAVGHME:KDAVRA
RGLLDRSFANMAEPPKVMYTNAGSVAENFLTMGMGFLQAVVFGCTGFRVTRAGVTFD
PVLCSGISRVSAGFYOENKLNFSFSDSVTEVTARAGVPWAPHLAEELWSPSRLS
L:PGHKVFFPRSGRAGIQMSPPKLPSSSSSEFFGRTFSDVRDPLQSLMWTT:GSSSPT
S:TVDPAGE"
```

BASE COUNT 1079 a 2146 c 2150 g 1086 t
ORIGIN

Query Match 58.4%; Score 1571.2; DB 6; Length 6461;
Best Local Similarity 89.6%; Pred. No. 4.5e-236;
Matches 1772; Conservative 0; Mismatches 38; Indels 168; Gaps 1;

```
Qy 7 CGTGGCGCGAGCCTGGCTGACCTGATCCTCGACAGTGGCCCGACCGCGGCGCGCGG 66
Db 731 CGGCGACGCGGAGCCTGGCTGACCTGATCCTCGACAGTGGCCCGACCGCGGCGCGG 790
Qy 67 TGGCGCAGATGTGGCCCGACGCGGCGCTCTTTCATCTCGACGCGCGGAGCAGC 126
Db 791 TGGCGCAGATGTGGCCCGACGCGGCGCTCTTTCATCTCGACGCGCGGAGCAGC 850
Qy 127 TGGCGCGCTGTGGGCGGCCCGGAGGCGCGCTGTCACAGACCCCTTCGAGGCGGAGCG 186
Db 851 TGGCGCGCTGTGGGCGGCCCGGAGGCGCGCTGTCACAGACCCCTTCGAGGCGGAGCG 910
Qy 187 GCGCGCGGTGTAGCGGGTGTGCTAGTAAAGCGCTGCTGCGCAGCGCCCTCTCTGCTGG 246
Db 911 GCGCGCGGTGTAGCGGGTGTGCTAGCAAGGCGCTGCTGCGCAGCGCCCTCTCTGCTGG 970
Qy 247 TGACCAACGCGCGCGCGCGCGCGGAGGCTGCAGCGCGCGCTGTGTTCCCGCAGTGGC 306
Db 971 TGACCAACGCGCGCGCGCGCGCGGAGGCTGCAGCGCGCGCTGTGTTCCCGCAGTGGC 1030
Qy 307 CCGAGGTGCGCGGCTTCTCCGACAAAGGACAAAGAAAGTATTTCACAAAGTTCCTCCGGG 366
Db 1031 CCGAGGTGCGCGGCTTCTCCGACAAAGGACAAAGAAAGTATTTCACAAAGTTCCTCCGGG 1090
```



```
repeat_region /db_xref="taxon:9606"
repeat_region /chromosome="11"
repeat_region /map="lip15.5"
repeat_region 711..975
repeat_region 2728..3022
repeat_region /rpt_family="Alu"
repeat_region 3363..3719
repeat_region /rpt_family="THE1"
repeat_region 5634..5909
repeat_region /rpt_family="Alu"
repeat_region complement(6720..6959)
repeat_region /rpt_family="Alu"
repeat_region 10560..10839
repeat_region /rpt_family="Alu"
repeat_region complement(11832..12455)
repeat_region /rpt_family="Alu"
repeat_region 12921..13405
repeat_region /rpt_family="Alu"
repeat_region 13718..14314
repeat_region /rpt_family="Alu"
repeat_region 14617..14779
repeat_region /rpt_family="Alu"
repeat_region 15336..15695
repeat_region /rpt_family="Alu"
repeat_region 15722..15994
repeat_region /rpt_family="Alu"
repeat_region complement(16896..16997)
repeat_region /rpt_family="Alu"
repeat_region complement(18996..19286)
repeat_region /rpt_family="Alu"
repeat_region complement(19458..19747)
repeat_region /rpt_family="Alu"
repeat_region complement(19777..20045)
repeat_region /rpt_family="Alu"
repeat_region complement(22885..23002)
repeat_region /rpt_family="MIR"
repeat_region 23109..23261
repeat_region /rpt_family="Alu"
repeat_region 23292..23481
repeat_region /rpt_family="MER46"
repeat_region complement(23877..23944)
repeat_region /rpt_family="MIR"
repeat_region complement(24695..24986)
repeat_region /rpt_family="Alu"
repeat_region complement(25429..26012)
repeat_region /rpt_family="Alu"
repeat_region complement(27380..27676)
repeat_region /rpt_family="Alu"
repeat_region complement(29422..29486)
repeat_region /rpt_family="MER42"
repeat_region complement(29514..29802)
repeat_region /rpt_family="Alu"
repeat_region 30797..31075
repeat_region /rpt_family="Alu"
repeat_region 31688..31792
repeat_region /rpt_family="Alu"
repeat_region 32125..32412
repeat_region /rpt_family="Alu"
repeat_region complement(38871..39186)
repeat_region /rpt_family="Alu"
repeat_region 42952..43094
repeat_region /rpt_family="MERS"
repeat_region 43106..43388
repeat_region /rpt_family="Alu"
repeat_region 43754..44216
repeat_region /rpt_family="MIR2B2"
repeat_region complement(44815..45136)
repeat_region /rpt_family="MER42"
repeat_region 45145..45484
repeat_region /rpt_family="Alu"
repeat_region 45740..45995
repeat_region /rpt_family="Alu"
```

```
repeat_region 46375..46646
repeat_region /rpt_family="Alu"
repeat_region 47145..47738
repeat_region /rpt_family="Alu"
repeat_region 48353..48626
repeat_region /rpt_family="Alu"
repeat_region 49567..49842
repeat_region /rpt_family="Alu"
repeat_region 49940..50211
repeat_region /rpt_family="Alu"
repeat_region complement(50615..50907)
repeat_region /rpt_family="Alu"
repeat_region complement(53001..53290)
repeat_region /rpt_family="Alu"
repeat_region complement(54161..54512)
repeat_region /rpt_family="L1"
repeat_region 54533..54823
repeat_region /rpt_family="Alu"
repeat_region 57899..58113
repeat_region /rpt_family="Alu"
repeat_region complement(58795..59044)
repeat_region /rpt_family="Alu"
repeat_region complement(59414..60023)
repeat_region /rpt_family="L1"
repeat_region complement(60177..60311)
repeat_region /rpt_family="Alu"
repeat_region complement(60770..61108)
repeat_region /rpt_family="L1"
repeat_region complement(61207..61557)
repeat_region 61670..61942
repeat_region /rpt_family="Alu"
repeat_region 61944..62436
repeat_region /rpt_family="MER41"
repeat_region complement(67807..68105)
repeat_region /rpt_family="Alu"
repeat_region 70961..71049
repeat_region /rpt_family="Alu"
repeat_region 71221..71523
repeat_region /rpt_family="Alu"
repeat_region complement(85033..85397)
repeat_region /rpt_family="MER4"
repeat_region complement(85414..85620)
repeat_region /rpt_family="Alu"
repeat_region complement(85701..85804)
repeat_region /rpt_family="MER4"
repeat_region complement(85895..86028)
repeat_region /rpt_family="MER4"
repeat_region 86180..86477
repeat_region /rpt_family="Alu"
repeat_region 86497..86780
repeat_region /rpt_family="Alu"
repeat_region 87125..87256
repeat_region /rpt_family="Alu"
repeat_region 87590..87893
repeat_region /rpt_family="Alu"
repeat_region 89103..89401
repeat_region /rpt_family="Alu"
repeat_region 89726..89837
repeat_region /rpt_family="Alu"
repeat_region 89888..90390
repeat_region /rpt_family="MER4"
repeat_region complement(91891..92187)
repeat_region /rpt_family="Alu"
repeat_region complement(93951..94017)
repeat_region /rpt_family="LTR12"
BASE COUNT 23552 a 24227 c 24425 g 22834 t
ORIGIN
```

Query Match 50.9%; Score 1369; DB 9; Length 95038;
Best Local Similarity 99.6%; Pred.No.1.2e-204;
Matches 1372; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 CGTGGCGCGAGCTGGCTGATCTGACCTGATCTGACAGTGCCTGAGCGCGCGCGCG 66
DB 64958 CGGCAACGCGAGCTGGCTGATCTGATCTGACAGTGCCTGAGCGCGCGCGCG 65017
QY 67 TGGCGAGAGTGTGGCGAGCGCGCGCTGCTCTCTCATCTCTGGAGCGCGCGAGCG 126
DB 65018 TGGCGAGAGTGTGGCGAGCGCGCGCTGCTCTCTCATCTCTGGAGCGCGCGAGCG 65077
QY 127 TGGCGAGTGTGGCGAGCGCGCGCTGCTCTCTCATCTCTGGAGCGCGCGAGCG 136
DB 65078 TGGCGAGTGTGGCGAGCGCGCGCTGCTCTCTCATCTCTGGAGCGCGCGAGCG 65137
QY 167 GCGCGCGGTGTAGCGGGTGTGCTGAGTGAAGCGCTGCTGCTGCTGCTGCTGCTG 246
DB 65138 GCGCGCGGTGTAGCGGGTGTGCTGAGTGAAGCGCTGCTGCTGCTGCTGCTGCTG 65197
QY 247 TGACACGCGCGCGCGCGCTGCGGGAGGTGACAGCGCGCTGCTGCTGCTGCTGCTG 306
DB 65198 TGACACGCGCGCGCGCGCTGCGGGAGGTGACAGCGCGCTGCTGCTGCTGCTGCTG 65257
QY 307 CCGAGGTGCGCGGTCTTCCAGAGGAGCAAGAGATTTCTACAGTCTTCCGGG 366
DB 65258 CCGAGGTGCGCGGTCTTCCAGAGGAGCAAGAGATTTCTACAGTCTTCCGGG 65317
QY 367 ATGAGAGAGCGCGCGCGCTTCCAGAGGAGCAAGAGATTTCTACAGTCTTCCGGG 426
DB 65318 ATGAGAGAGCGCGCGCGCTTCCAGAGGAGCAAGAGATTTCTACAGTCTTCCGGG 65377
QY 427 TGTGCTTGTGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
DB 65378 TGTGCTTGTGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65437
QY 487 TCGTGTGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546
DB 65438 TCGTGTGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 65497
QY 547 CCAGCGTCTGAGCTGCGCTGCGTACGCGAGCGCGCGCGCTGCTGCTGCTGCTGCTG 606
DB 65498 CCAGCGTCTGAGCTGCGCTGCGTACGCGAGCGCGCGCGCTGCTGCTGCTGCTGCTG 65557
QY 607 ATCTGTGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
DB 65558 ATCTGTGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 65617
QY 667 AACTGGAGCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
DB 65618 AACTGGAGCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 65677
QY 727 AGCTGCGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
DB 65678 AGCTGCGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65737
QY 787 AGTTCTTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846
DB 65738 AGTTCTTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65797
QY 847 GCGCGTGTGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
DB 65798 GCGCGTGTGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65857
QY 907 GCGCGTGTGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
DB 65858 GCGCGTGTGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65917
QY 967 GCTGATGTTTACAGCGTGTGAGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
DB 65918 GCTGATGTTTACAGCGTGTGAGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65977
QY 1027 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086
DB 65978 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 66037
QY 1087 AAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1146

DB 66038 AAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 66097
QY 1147 GCTGTACGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1206
DB 66098 GCTGTACGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 66157
QY 1207 TGGCGCTGACGAGTGGCTTCTGCGCATGGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTG 1266
DB 66158 TGGCGCTGACGAGTGGCTTCTGCGCATGGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTG 66217
QY 1267 GGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326
DB 66218 GGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 66277
QY 1327 AGAAGAAGAAGAGGCTGCGGAGCGGCTTCCAGGCGAGCCTGCGTGGTGGCGGAGT 1383
DB 66278 AGAAGAAGAAGAGGCTGCGGAGCGGCTTCCAGGCGAGCCTGCGTGGTGGCGGAGT 66334

RESULT 11
AC136475
LOCUS
DEFINITION
AC136475
ACCESSION
AC136475
VERSION
AC136475.6
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 143835)
AUTHORS
Biren, B., Nussbaum, C. and Lander, E.
TITLE
Homo sapiens chromosome 11, clone RP11-326C3
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 143835)
AUTHORS
Biren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,
Mathews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, N.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N.,
Stojatovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, N.
Direct Submission
TITLE
Submitted (02-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL
3 (bases 1 to 143835)
REFERENCE
3 (bases 1 to 143835)
AUTHORS
Biren, B., Nussbaum, C., Lander, E., Aboueleil, A., Allen, N.,
Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collamore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,
Mathews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, N.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N.,
Stojatovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, N.
Direct Submission
TITLE
Submitted (02-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
JOURNAL
3 (bases 1 to 143835)
REFERENCE
3 (bases 1 to 143835)
AUTHORS


```
QY 1118 G---CCCAACTACCCACTGGAAGTGTCTGTACTGCTGTACGAGACGAGGAGGAGCGTT 1174
DB |||||
QY 893 GAGACTCAATTTTGACTGGAGCTGCTGTACTGCTGTATGAGACACAGGAGGAGGATTT 952
DB |||||
QY 1175 TGTGCGCAAGCCCTGTGCGCTGTTCCCGAGCTGGCGCTGCAGGAGTGTCTTCTCCG 1234
DB |||||
QY 953 TTTTCGACAGGCTCTCAGACGCTTCCAGAGATAGTACTGGAGCGAGTCAAGTTGACCCG 1012
DB |||||
QY 1235 CATGACGCTGGCTGTCTGAGCTACTGCTGAGGTGCTGCCCTGCTGCAGAGGCACTGCG 1294
DB |||||
QY 1013 CATGACCTTGAAGTCTGAATTAATCTGCTGCAAGTCTGCCAGATGCCAGGCTCTGAG 1072
DB |||||
QY 1295 GCTGATCAGCTCAGATTTGTTGCTGCGCAGAGGAAGGAAGAGAGCCCTGGGGAGGCG 1354
DB |||||
QY 1073 ACTGGTGAAGCTGGGCTGTGTCGGCAAGAGGAAGAAGAAGA-----AGAAGAG 1126
DB |||||
QY 1355 GCTCCAGGCCAGCTGGGTGGCGGAGTTCTCAAGSCACCAAAACAACTGCCAGCCTC 1414
DB |||||
QY 1127 CTGTGTGAAGCGCTGAAGGGC-----TCTCAAGCACCAGAAACAAACCCAGTCTC 1180
DB |||||
QY 1415 CTTCTTTTCATCACTCTTTTCAGGCAATGACTGACCCACTGTGCCATCTGACGACGCTCAC 1474
DB |||||
QY 1181 CTTGCTGCTGCCACTCTGTGACAACTGACTACCCCGAAATGTCATCTGAGTGTCTGAT 1240
DB |||||
QY 1475 GCTGTCCCACTGCAAACTCTCTGACCGCTCTGCCAGACCTTTCTGAGGCCCTGAGGC 1534
DB |||||
QY 1241 CTGTGTCACTGACAGCTCTCTGATGCGAGTTTGGCGAGACCTTTCCGAGGCCCTGAAGT 1300
DB |||||
QY 1535 AGCCCGGCACTGACGAGCTGGGCTCTCTCCACAAACAGGCTCACTGAGGCGGAGCTGCG 1594
DB |||||
QY 1301 AGCTCTGCCCTAAGGAGTTGGGCTCTCCAGAGCGGCTCCACACACAGAGCTGCG 1360
DB |||||
QY 1595 TATGCTAGTAGGCGCTAGCTGCGCGAGTGTCAGGCTGACAGCGTCAAGGTACAGCT 1654
DB |||||
QY 1361 TTAGTGTGCAAGGCTGGCTGTGCGCAAGTGGCGAGCTTCCGAGGCCCTGAAGT 1420
DB |||||
QY 1655 GCTGACCCCGAGCGAGGCTCCAGTACTGCTGGGTATGCTCGCGAGAGCCCGGCT 1714
DB |||||
QY 1421 GCTTACCTTCAAGAGTGATCAATTCTTGTGATCTGCTCAGCAGAGCCAGTCTCT 1480
DB |||||
QY 1715 GACCACTTGGATCTGAGCGCTGCCAACTGCCCGCCCGCCCTACCTTGTGTGC 1774
DB |||||
QY 1481 GACCACTTGGACCTCAGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1540
DB |||||
QY 1775 AGTCTGAGCAGCAGGAGTGGGCTGAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1834
DB |||||
QY 1541 AGCCCTGAAGCAGCCCTAAATGCTCCCTAAAGACCTCTGCTGCTGCTGCTGCTGCTGCT 1600
DB |||||
QY 1835 CGAGAGTCACTACAGAGCTTCAGGCTGTGAAGAGCAAGCCGATCTGGTCACTAC 1894
DB |||||
QY 1601 TGAGAACTCACTGAGGATCTTCAAGCTGTGAAGACATCAAGCCAGATCTGCTCATCAT 1660
DB |||||
QY 1895 ACACCGAGCGCTGGAGCGGCAACCAACACTCCCAAGGAATCTATCTGACCTTCTGAGG 1954
DB |||||
QY 1661 ATATTCAAAATAAGGCACACCTCAACCGGCTCGGAAGG-GTSAGGCACTGCTTCTGAAG 1719
DB |||||
QY 1955 CTCTGCTGCGCAGAGCAGGCTGGAAGACCT---AGTCAAGTCCCTCTGAGAGACGG 2011
DB |||||
QY 1720 CTGCGAGTCTCAAGACACAGTGAAGATGCTGTAGCCCAAGATCTCTGTAATGAGCT 1779
DB |||||
QY 2012 CCATTCCAAAGGCGAGGAGATATTGCTCTCGGCTTTTGGGAACTTTTGAAGCGGAGG 2071
DB |||||
QY 1780 CCCATTCAAAAGCAAGAGATGCCAGCTTCTCCCTCCAGGCAAGTCTTCAAGGCCAG 1839
DB |||||
QY 2072 CCGCAGACAGGATGTGGAGGCCGAGACAGCGACCTTCCCGCTCCAGGACAGGCCCA 2131
DB |||||
QY 1840 AAG--GGCAAGCAAAAGAC--CAGGTAGATACAGGCGCTTGTAGTCTCTGATAGACCTG 1897
DB |||||
QY 2132 GGAAGTGGCCCTCTCTCCA--ACTGCGGTTACCCCTTCTCCCGCCAGCCCACTACTCC 2191
DB |||||
QY 1898 GACAGTGGCTGGCTGGGA--ACTCTGCGGTTGGTCTTCCCTCAGTCCCAACACAGCAAT 1957
DB |||||
QY 2192 ACCACCTTCTCT 2203
```

```
DB 1959 GCCTTCCCTTCT 1969 || |||||
RESULT 15
AP006286 175416 bp DNA linear HTG 26-MAR-2003
LOCUS Homo sapiens chromosome 11 clone RP13-25N22 map 1-p, WORKING DRAFT
DEFINITION SEQUENCE, 22 unordered pieces.
ACCESSION AP006286
VERSION AP006286.1 GI:29243345
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens genomic DNA of 11p
JOURNAL Published Only in Database (2003)
REFERENCE 2 (bases 1 to 175416)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suhiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gscc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT ----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gscc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP13-25N22
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 160074 bases at least Q40
Consensus quality: 167722 bases at least Q30
Consensus quality: 171662 bases at least Q20
Insert size: 17316; sum-of-contents
Quality coverage: 7.92x in Q20 bases; sum-of-contents
-----
NOTE: This is a 'working draft' sequence. It currently consists of
22 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 33131 contig of 33131 bp in length
33232 54098 contig of 20867 bp in length
54199 72037 contig of 7839 bp in length
72138 83277 contig of 11240 bp in length
83378 98599 contig of 15222 bp in length
98700 105513 contig of 6814 bp in length
105614 116229 contig of 10616 bp in length
116330 121964 contig of 5635 bp in length
122065 127627 contig of 5563 bp in length
127728 131971 contig of 4244 bp in length
132072 137506 contig of 5435 bp in length
137607 143829 contig of 6223 bp in length
143930 148539 contig of 4610 bp in length
148640 153148 contig of 4509 bp in length
153249 157509 contig of 4261 bp in length
157610 160561 contig of 2952 bp in length
160662 164778 contig of 4117 bp in length
164879 168269 contig of 3391 bp in length
```


Qy	2291	ACCTACCTGGGGCAGAAATACCCCTTCACTGCTGATCCCATCTGCAGTGTGGCCCAAC	2350
Db	102344	ACCTACCTGGGGCAGAAATACCCCTTCACTGCTGATCCCATCTGCAGTGTGGCCCAAC	102403
Qy	2351	AGCCCCCAGAACTATGCCCCACATAGACTGGAGGTAGGCAGTTCACCGTCCCTCCCTGTTA	2410
Db	102404	AGCCCCCAGAACTATGCCCCACATAGACTGGAGGTAGGCAGTTCACCGTCCCTCCCTGTTA	102463
Qy	2411	GGAAATGAGACCATCCCTGAGCTATGGCCCAAGCCGACAGGGCTCCAGTGTGAGATCT	2470
Db	102464	GGAAATGAGACCATCCCTGAGCTATGGCCCAAGCCGACAGGGCTCCAGTGTGAGATCT	102523
Qy	2471	TTGGGAAGGGAGACTAGGGCAGGTGGAGACAGCGAGAACCCCGTGTGGGTGGGAAGC	2530
Db	102524	TTGGGAAGGGAGACTAGGGCAGGTGGAGACAGCGAGAACCCCGTGTGGGTGGGAAGC	102583
Qy	2531	ATGACCACATGGTGGGTGAGCAGCCGCCCATGCACTGACGGTAAATTCGCCCTGTGGACTCA	2590
Db	102584	ATGACCACATGGTGGGTGAGCAGCCGCCCATGCACTGACGGTAAATTCGCCCTGTGGACTCA	102643
Qy	2591	TTTCTGTGGTCTCTATTACACTGGCCAGGGTGGTACAAATACAGGTGGGTGCTCACA	2650
Db	102644	TTTCTGTGGTCTCTATTACACTGGCCAGGGTGGTACAAATACAGGTGGGTGCTCACA	102703

Search completed: October 3, 2003, 11:59:29
Job time : 9833 secs

GenCcre version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 17:35:13 ; Search time 86 Seconds
(without alignments)
153.536 Million cell updates/sec

Title: US-10-029-347-2

Perfect score: 3218

Sequence: 1 MIAQPORLLILDGADLPA.....ITHPALDGHQPQKELISTF 625

Scoring table: BLOSUM62

Gapop 10.0 , Gapext: 0.5

Searched: 1:07863 seqs, 15876573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	3218	100.0	625	23	ABG78454	Human leucine-rich
2	3218	100.0	651	22	AAE04546	Human G-Protein co
3	3144	97.7	676	23	ABU085075	Human NOV18a prote
4	2882	89.6	1551	22	AAU00023	Human activated T-
5	2882	89.6	1851	23	AAO17860	Pyrim cor-ain conta
6	2239	69.6	597	22	ABB11735	Human vasopressin
7	1578	49.0	481	24	ABR33389	AI/AVPv2 receptor
8	1220	34.8	321	23	ABB90173	Human polypeptide
9	1056.5	32.8	329	22	AAH89960	Human immune/haema

10	825.5	25.7	1061	23	AAO15590	Human PYRIN-8 prot
11	825	25.6	1099	23	AAO17857	Pyrim domain conta
12	800.5	24.9	1034	22	AAE07514	Human PYRIN-1 prote
13	800.5	24.9	1034	24	ABU08503	Human PYRIN-1 prote
14	762	23.7	1429	22	AAE62571	Human CARD-7 polyp
15	762	23.7	1429	23	ABG97969	Human leucine rich
16	762	23.7	1429	23	ABG78472	Leucine-rich repea
17	762	23.7	1429	23	ABB77916	Human leucine-rich
18	762	23.7	1429	24	ABG71631	Human caspase recr
19	762	23.7	1429	24	ABG71633	Human caspase recr
20	762	23.7	1454	22	AAE06758	Human NAC beta iso
21	762	23.7	1473	22	AAE06758	Human G-protein co
22	762	23.7	1473	22	AAE06758	Human NB-ARC and C
23	762	23.7	1473	22	AAE06758	Human NAC beta iso
24	760	23.6	1429	23	ABG78455	Human caspase recr
25	760	23.6	1429	23	AAO17855	Pyrim domain conta
26	716	22.2	1424	22	AAE06758	Human NAC gamma or
27	712.5	22.1	1397	22	AAE06758	Human NB-ARC and C
28	712.5	22.1	1442	22	AAE06758	Human NB-ARC and C
29	691	21.5	920	23	ABG78454	Human MDDT-13 prot
30	660	20.5	952	22	ABG03924	Novel human diagno
31	650	20.2	980	23	ABG98244	Human CDD SEQ ID
32	650	20.2	980	23	AAH50329	Human nucleotide b
33	641	19.9	449	23	ABG77910	Amino acid sequenc
34	635	19.7	1022	23	AAO17862	Pyrim domain conta
35	634	19.7	1049	22	AAE065894	Amino acid sequenc
36	619.5	19.3	1033	22	AAE07512	Human nucleotide b
37	619.5	19.3	1033	23	ABG78456	Human nucleotide b
38	619.5	19.3	1033	24	ABU08502	Human pyrim domain
39	616.5	19.2	1062	22	AAE065895	Amino acid sequenc
40	616.5	19.2	1062	23	AAO17856	Pyrim domain conta
41	598.5	18.6	858	23	AAH50328	Human nucleotide b
42	598.5	18.6	919	23	ABG43483	Human secreted pro
43	598.5	18.6	919	24	ABU19319	NOVX related prote
44	598.5	18.6	983	23	AAO17870	Pyrim domain conta
45	598.5	18.6	994	23	AAO15593	Human PYRIN-3 prot

ALIGNMENTS

RESULT 1
ABG78454
ID ABG78454 standard; Protein; 625 AA.
XX ABG78454;
XX
XX 15-NOV-2002 (first entry)
XX Human leucine-rich repeat small intestine I (HLRSL1).
XX Human; human leucine-rich repeat, small intestine I; HLRSL1; asthma;
XX proliferative disorder; gastrointestinal disorder; renal disorder;
XX neural disorder; reproductive disorder; calcium regulation; apoptosis;
XX immune system; anaemia; human immune deficiency virus; HIV; cancer;
XX blood coagulation disorder; autoimmune disorder; allergic reaction;
XX inflammatory condition; cardiovascular disorder; ischaemia;
XX neurological disorder; infectious disease; cytokine production;
XX expressed sequence tag; EST.
XX Homo sapiens.
XX WQ200261086-A2.
XX
XX 08-AUG-2002.
XX
XX 20-DEC-2001; 2001WO-US49739.
XX
XX 22-DEC-2000; 2000US-25774P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Feder J, Ramanathan C, Mintier G;

XX
DR WPI: 2002-6.9252/66.
XX N-PSDB: ABS63485.
XX
PT New isolated nucleic acid molecules encoding HLRRSII polypeptides, or
PT their fragments and homologues, useful for preventing, treating and
PT ameliorating medical conditions, e.g. proliferative, gastrointestinal,
PT or renal disorders.
XX
XX Claim 20; Figure 2; 336pp; English.
XX
XX The invention relates to isolated nucleic acid molecules (I) encoding
CC human leucine-rich repeat small intestine I (HLRRSII) polypeptides.
CC The nucleic acid molecules and polypeptides are useful for preventing,
CC treating and ameliorating medical conditions, such as proliferative,
CC gastrointestinal, renal, neural, or reproductive disorders; or disorders
CC related to aberrant calcium regulation or apoptosis modulation, either
CC directly or indirectly. They are also useful for treating, preventing
CC and/or diagnosing diseases, disorders and/or conditions of: immune system
CC by activating or inhibiting the proliferation, differentiation, or
CC mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
CC anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
CC virus (HIV) infection, HTLV-BIV infection; blood coagulation disorders,
CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
CC myasthenia gravis; asthma or allergic reactions; inflammatory
CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
CC ischaemia; aneurysms; neurological disorders, e.g. Alzheimer's disease,
CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
CC pneumonia, or viral, bacterial, and fungal infections. The HLRRSII
CC polypeptides are useful for modulating cytokine production, antigen
CC presentation, or other processes such as boosting immune responses.
CC AS578454-ABG78474 represent HLRRSII amino acid sequences and related
CC amino acid sequences of the invention.
XX
SQ Sequence 625 AA;

Query Match 100.0%; Score 3218; DB 23; Length 625;
Best Local Similarity 100.0%; Pred. No. 5.3e-314;
Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIAQFORLLFILDGADELPAAGPEAPCTDPPEAASGARVLGGLLSKALLPTALLVVT 60
DQ 1 YLAQFORLLFILDGADELPAAGPEAPCTDPPEAASGARVLGGLLSKALLPTALLVVT 60
QY 61 RAAAPGRJQRLCSPOCEVRGFSKDKKKYFYKFFDERRAERAYRFVKNETLFAJCF 120
DQ 61 RAAAPGRJQRLCSPOCEVRGFSKDKKKYFYKFFDERRAERAYRFVKNETLFAJCF 120
QY 121 VPFVCWIVCTVLRQLEGRDLRSKTTTTSVLLFTSVLSSAPVADGPPLOGLANLC 180
DQ 121 VPFVCWIVCTVLRQLEGRDLRSKTTTTSVLLFTSVLSSAPVADGPPLOGLANLC 180
QY 181 RLAREGVLRRAQFAEKELEGLRGSKVQTLF-SKKELPGVLETEVTFIDQSFEFL 240
DQ 181 RLAREGVLRRAQFAEKELEGLRGSKVQTLF-SKKELPGVLETEVTFIDQSFEFL 240
QY 241 AALSYLEDDGVPRTAAGGVQTLRGDAQPHSHVLVTRFLFGLLSAERMRDIERHFGCM 300
DQ 241 AALSYLEDDGVPRTAAGGVQTLRGDAQPHSHVLVTRFLFGLLSAERMRDIERHFGCM 300
QY 301 VSERVKQBALRWQGGGCGPQVAPETEGAGKLEDEEPEEEEGEPNYPLELYCLY 360
DQ 301 VSERVKQBALRWQGGGCGPQVAPETEGAGKLEDEEPEEEEGEPNYPLELYCLY 360
QY 361 ETQEDAFVRQALCRPELALORVFCRMDVAVLSYVRCRCCPAGQALRLISRLVAAGEKK 420
DQ 361 ETQEDAFVRQALCRPELALORVFCRMDVAVLSYVRCRCCPAGQALRLISRLVAAGEKK 420
QY 421 KKSGLKRLQASLGGSSGGTQKQPASLLHPLFOAMTDPPLCHLSLTLSHCKLPDVCRD 480
DQ 421 KKSGLKRLQASLGGSSGGTQKQPASLLHPLFOAMTDPPLCHLSLTLSHCKLPDVCRD 480

QY 481 LSEALRAAPALTELGLLHRLSEAGLRMLSEGLAWPQCRVQTVRVOLPDPQSGLOYLVGM 540
DQ 481 LSEALRAAPALTELGLLHRLSEAGLRMLSEGLAWPQCRVQTVRVOLPDPQSGLOYLVGM 540
QY 541 LRQSPALTTLDLSGGQQLPAPMVTYLCVQLHOGCGLOTLASVELSQSLQELQAVKRA 600
DQ 541 LRQSPALTTLDLSGGQQLPAPMVTYLCVQLHOGCGLOTLASVELSQSLQELQAVKRA 600
QY 601 KPDLVITHPALDGHPPPKELISTF 625
DQ 601 KPDLVITHPALDGHPPPKELISTF 625

RESULT 2
ID AAE04546 standard; Protein; 891 AA.
XX AAE04546;
XX
XX 04-SEP-2001 (first entry)
XX Human G-protein coupled receptor-2 (GCRC-2) protein.
XX Human: G-protein coupled receptor-2; GCRC-2; gene therapy; cirrhosis;
XX transgenic animal; proliferative disorder; actinic keratosis; hepatitis
XX nephrotropic; cancer; breast; bladder; bone marrow; brain; uterus; stroke;
XX leukaemia; adenocarcinoma; lymphoma; melanoma; myeloma; epilepsy; stroke;
XX neurological disorder; Alzheimer's disease; Parkinson's disease; nausea;
XX Huntington's disease; multiple sclerosis; dementia; angina pectoris;
XX central nervous system disorder; cardiovascular disorder; hypertension;
XX atherosclerosis; congestive heart failure; gastrointestinal disorder;
XX dysphagia; peptic oesophagitis; spasm; gastritis; anorexia; pyrosis;
XX pancreatitis; Crohn's disease; diarrhoea; autoimmune disorder; anaemia;
XX inflammatory disorder; Acquired Immune Deficiency Syndrome; AIDS;
XX Addison's diseases; allergy; asthma; diabetes mellitus; antithyroid;
XX atopic dermatitis; glomerulonephritis; Grave's disease; psoriasis;
XX rheumatoid arthritis; ulcerative colitis; osteoporosis; antifungal;
XX metabolic disorder; obesity; noctropic; protozoacide; virucide.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Binding-site 202..209 /label= P loop
XX /note= "ATP/GTP binding site"
XX W0200142288-A2.
XX 14-JUN-2001.
XX
XX 07-DEC-2000; 2000MO-US33382.
XX
XX 10-DEC-1999; 99US-0172852.
XX 22-DEC-1999; 99US-0171732.
XX 14-JAN-2000; 2000US-0176148.
XX 21-JAN-2000; 2000US-0177331.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX Barford N, Baughn MR, Au-Young J, Yang J, Lu DAN, Reddy R;
XX WPI: 2001-381635/40.
XX N-PSDB: AAD08636.
XX
XX New human G-protein coupled receptor polypeptides for diagnosing,
XX preventing, and treating cell proliferative, neurological,
XX cardiovascular, gastrointestinal, autoimmune and metabolic disorders -
XX Claim 1; Page 129-131; 175pp; English.
XX
XX The present sequence is human G-protein coupled receptor-2 (GCRC-2)
XX protein. GCRC-2 is useful in somatic or germline gene therapy to correct
XX a genetic deficiency, to express a conditionally lethal gene product and

CC to express a protein which affords protection against intracellular
 CC parasites and also for diagnosis of disorders associated with expression
 CC of CREC. CREC is also useful for generating hybridisation probes useful
 CC in mapping the naturally occurring genomic sequences and to create
 CC knockin humanised animals (pigs) or transgenic animals (mice or rats) to
 CC model human diseases. CREC is used to diagnose, prevent and treat
 CC proliferative disorders (actinic keratosis, arteriosclerosis, cirrhosis,
 CC hepatitis and cancer); cancer (breast, bladder, bone marrow, brain,
 CC uterus cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma)
 CC neurological disorders (epilepsy, stroke, Alzheimer's, Huntington's,
 CC Parkinson's disease, multiple sclerosis, dementia and other central
 CC nervous system disorders); cardiovascular disorders (angina pectoris,
 CC hypertension, atherosclerosis, congestive heart failure);
 CC gastrointestinal disorders (dysphagia, peptic oesophagitis, oesophageal
 CC spasm, gastritis, gastric carcinoma, anorexia, nausea, abdominal angina,
 CC pyrosis, pancreatitis, Crohn's disease, diarrhoea); autoimmune/
 CC inflammatory disorders (acquired immunodeficiency syndrome (AIDS),
 CC Addison's disease, allergies, anaemia, asthma, diabetes mellitus, atopic
 CC dermatitis, glomerulonephritis, Grave's disease, osteoarthritis,
 CC psoriasis, rheumatoid arthritis, ulcerative colitis, bacterial, fungal,
 CC parasitic, protozoal and helminthic infections) and metabolic disorders
 CC (obesity, osteoporosis, viral infections).

XX
 SQ Sequence 891 AA;
 Query Match 100.0%; Score 3218; DB 22; Length 891;
 Best Local Similarity 100.0%; Pred No. 9,2e-314;
 Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLAQORLLFLIDGADLPALGPEAPCTDPFEAASGARVLGGLSKALLPTALLVTT 60
 DB 267 MLAQORLLFLIDGADLPALGPEAPCTDPFEAASGARVLGGLSKALLPTALLVTT 326
 QY 61 RAAAPRLGRLCSPCAVRFSKXKKYKFRDERRAERYRVRKNETLFLALCF 120
 DB 327 RAAAPRLGRLCSPCAVRFSKXKKYKFRDERRAERYRVRKNETLFLALCF 386
 QY 121 VFVCHVICTVLRQQLLEGRD-SRTSKTTTSVYLLFTSVLSSAPVADGPRQGLRLNC 180
 DB 387 VFVCHVICTVLRQQLLEGRD-SRTSKTTTSVYLLFTSVLSSAPVADGPRQGLRLNC 446
 QY 181 RLAREGLVGRRAQFAEKELEQ-ELRSGSKVQTLFLSKKELPGVJETEVYQFIDQSFQEF 240
 DB 447 RLAREGLVGRRAQFAEKELEQ-ELRSGSKVQTLFLSKKELPGVJETEVYQFIDQSFQEF 506
 QY 241 AALSYLEGGVPRTRAGGVTLRGDAQPHSHLV-TTRFLFGLLSAERMDIERHFCQM 300
 DB 507 AALSYLEGGVPRTRAGGVTLRGDAQPHSHLV-TTRFLFGLLSAERMDIERHFCQM 566
 QY 301 VSERVVKQEAALRWQGGGCGPVAPVETEGAKGLEDTPEPEEEEGEENYPLELLYCLY 360
 DB 567 VSERVVKQEAALRWQGGGCGPVAPVETEGAKGLEDTPEPEEEEGEENYPLELLYCLY 626
 QY 361 ETQEDAFVQALCRFPPELALQVRFCRMDVAVLSYVRCPCAGQALRLISRLVAQEKX 420
 DB 627 ETQEDAFVQALCRFPPELALQVRFCRMDVAVLSYVRCPCAGQALRLISRLVAQEKX 686
 QY 421 KKSGLGRLOASLGSSGSGTQKQ-PASLALHPLFOAMTDP-CHLSLITLHSHCKLPDAVCRD 480
 DB 687 KKSGLGRLOASLGSSGSGTQKQ-PASLALHPLFOAMTDP-CHLSLITLHSHCKLPDAVCRD 746
 QY 481 LSEALRAAPALTTELGLHNLRLSEAGLRMLSEG-LAMPQCRVQTVRVQLPDPGRGLQYLVGX 540
 DB 747 LSEALRAAPALTTELGLHNLRLSEAGLRMLSEG-LAMPQCRVQTVRVQLPDPGRGLQYLVGX 806
 QY 541 LROSPALTTLDLSSGCGPAPMTYILCAV-CHQCGGLQTLTSLASVELSOSLOELQAVKRA 600
 DB 807 LROSPALTTLDLSSGCGPAPMTYILCAV-CHQCGGLQTLTSLASVELSOSLOELQAVKRA 866
 QY 601 KPDLVITHPALDGHGHPKPELSTF 625
 DB 867 KPDLVITHPALDGHGHPKPELSTF 891

RESULT 3

ABJ65075
 ID ABU65075 standard; Protein; 676 AA.
 XX AC ABU65075;
 XX DT 20-MAY-2003 (first entry)
 XX DE Human NOV18a protein.
 XX KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KW human.
 XX OS Homo sapiens.
 XX PN WC200272757-A2.
 XX PD 19-SEP-2002.
 XX PF 08-MAR-2002; 2002MO-US06908.
 XX PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276300P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 22-MAR-2001; 2001US-277239P.
 PR 26-MAR-2001; 2001US-277321P.
 PR 21-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278936P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-277338P.
 PR 30-MAR-2001; 2001US-279995P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 04-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291039P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.

PR 18-OCT-2001; 2001US-33380P.
 PR 31-OCT-2001; 2001US-33530P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 21-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
 PI Zehrusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Patturajan M, Gangoli E, Vernet CAM, Guo X, Tchernev V;
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
 PI Anderson D, Spaderina SK, Catterton E, Burgess C, Leite M, Zhong H;
 PI Alsobrock JP, Lepley DM, Rieger DK;
 XX
 PI WPI; 2002-723332/78.
 DR N-PSDB; ABX97042.
 DR
 XX
 PT NOVX polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOVX expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma -
 XX
 PS Claim 1; Page 149; 1103pp; English.
 XX
 CC This invention describes novel human NOVX polypeptides which have
 CC cytostatic, cardiac, antiatherosclerotic, antiautomatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOVX
 CC proteins or nucleic acid molecules or NOVX antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOVX
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABU65041-ABU65219 represent
 CC the NOVX polypeptides encoded by ABX97008-ABX97185.
 XX
 SQ Sequence 676 AA;

Query Match 97.7%; Score 3144; DB 23; Length 676;
 Best Local Similarity 98.4%; Pred. No. 1.7e-306;
 Matches 615; Conservative 6; Mismatches 4; Indels 6; Gaps 1;

QY 1 MLAGPQRLFLDGADELPAJGPEAAPCTDPEAASGARVLGGLSKALLPTALLVTT 60
 DB 58 MLAGPQRLFLDGADELPAJGPEAAPCTDPEAASGARVLGGLSKALLPTALLVTT 117
 QY 61 RAAAPGLQRLCSQCAEVGSDKDKKYYFKFFDERRAERAIRFKENETLFAJCF 120
 DB 118 RAAAPGLQRLCSQCAEVGSDKDKKYYFKFFDERRAERAIRFKENETLFAJCF 177
 QY 121 VPFCVMTVCTVLRQELGRLSHTSTTTTSVLLFTSVLSAPVADGPRLOGDLRLNC 180
 DB 178 VPFCVMTVCTVLRQELGRLSHTSTTTTSVLLFTSVLSAPVADGPRLOGDLRLNC 237
 QY 181 RLAREGVGLRRAPAEKELELRLGSKVOTLFLSKKELFGVLETVTYQFIDQSFQEF 240
 DB 238 RLAREGVGLRRAPAEKELELRLGSKVOTLFLSKKELFGVLETVTYQFIDQSFQEF 297
 QY 241 AALSYLEDDGVPRTAAGGVJTLRGDAQPHSHLVLTTRFLGLSAERMEDIERHFGCM 300
 DB 298 AALSYLEDDGVPRTAAGGVJTLRGDAQPHSHLVLTTRFLGLSAERMEDIERHFGCM 357
 QY 301 VSERVKQEARLWVGQGGCPGVAPEVTEGAKGLEDTEEPFESEEEBEPNYPLELYCLY 360
 DB 358 VSERVKQEARLWVGQGGCPGVAPEVTEGAKGLEDTEEPFESEEEBEPNYPLELYCLY 417

QY 351 ETOEDAFVROALCRFPPELALQRFVCRMDVAVLSYCVRCPCAGCALRLISCLVAAQEK 420
 DB 418 ETOEDAFVROALGRFPPELALQRFVCRMDVAVLSYCVRCPCAGCALRLISCLVAAQEK 477
 QY 421 KKSJGKRLQASLGSGSGTTKQLPASLLHPLFOAMTDPLCHLSLTLSSHCKLPDVCRD 480
 DB 478 KKSJGKRLQASL-----GTTKQLPASLLHPLFOAMTDPLCHLSLTLSSHCKLPDVCRD 531
 QY 481 LSEALRAAPALTEGLLHNLRLSEAGLRLSEGGLAWPQCRVQTVRVQLPDPGRGLQYLVGM 540
 DB 532 LSEALRAAPALTEGLLHNLRLSEAGLRLSEGGLAWPQCRVQTVRVQLPDPGRGLQYLVGM 591
 QY 541 LRQSPALTTDLSCQLPAPMVTYLCVAVLQHOGLQTLTLASVELSEQSLCELQAVKRA 600
 DB 592 LRQSPALTTDLSCQLPAPMVTYLCVAVLQHOGLQTLTLASVELSEQSLCELQAVKRA 651
 QY 601 KPDLVITHPALDGHPPPKELISTF 625
 DB 652 KPDLVITHPALDGHPPPKELISTF 676
 RESULT 4
 AAU00023
 ID AAU00023 standard; Protein; 1851 AA.
 XX
 AC AAU00023;
 XX
 DT 10-MAY-2001 (first entry)
 XX
 DE Human activated T-lymphocyte associated sequence 2, ATLAS-2.
 XX
 KW Human; activated T-lymphocyte associated sequence 2; ATLAS-2; antibody;
 KW cytokine receptor; autoimmune disorder; immune disorder; cancer;
 KW T-lymphocyte-associated disorder; cell proliferation disorder; tumour;
 KW cell differentiation disorder; immune deficiency disorder; malignancy;
 KW viral infection; bacterial infection; fungal infection; metabolism;
 KW chromosome 11p15.5.
 XX
 OS Homo sapiens.
 XX
 FN WO200114564-A2.
 XX
 PD 01-MAR-2001.
 XX
 PF 15-AUG-2000; 2000WO-US22699.
 XX
 PR 20-AUG-1999; 99US-0150105.
 PR 28-APR-2000; 2000US-0560101.
 PR 28-APR-2000; 2000US-0560365.
 PR 28-APR-2000; 2000US-0560948.
 PR 28-APR-2000; 2000US-0561533.
 XX
 PA (CURA-) CURAGEN CORP.
 PA (B:OJ) BIOGEN INC.
 XX
 PI Peyman JA, Green CD, Hsu A, Browning JA, Carulli C;
 XX
 DR WPI; 2301-218453/22.
 DR N-PSDB; AAS00033.
 XX
 PT New isolated activated T lymphocyte associated sequences for treating
 PT or preventing immune system associated disorders such as autoimmune
 PT disorder, immune disorder, and T-lymphocyte-associated disorder -
 XX
 PS Claim 14; Fig 2; 114pp; English.
 XX
 CC The sequence represents human activated T-lymphocyte associated sequence
 CC 2, ATLAS-2. ATLAS-2 is related by homology to cytokine receptors and its
 CC gene is located on chromosome 11p15.5. ATLAS proteins, polynucleotides
 CC and antibodies are useful for treating/preventing conditions associated
 CC with an autoimmune disorder, immune disorder, T-lymphocyte-associated
 CC disorder, cell-proliferation disorder, cell differentiation disorder,
 CC and immune deficiency disorder and for screening for a modulator of

CC activity or of latency or predisposition to an immune disorder.
 CC ATLAS proteins, polynucleotides and antibodies are useful in therapeutic
 CC or prophylactic treatment of diseases associated with cell proliferation
 CC (e.g. cancers, malignancies and tumours). The polynucleotides are useful
 CC in gene therapy, to detect ATLAS mRNA or a genetic lesion in an ATLAS
 CC gene, to modulate ATLAS activity, to screen drugs or compounds that
 CC modulate ATLAS activity or expression and to treat disorders
 CC characterised by insufficient or excessive production of ATLAS protein or
 CC production of ATLAS protein forms that have decreased or aberrant
 CC activity compared to ATLAS wild type protein, and in tissue typing to
 CC identify individuals. The antibodies are useful for localisation/
 CC quantitation, isolation and detection of ATLAS and to monitor protein
 CC levels in tissue. ATLAS is useful for treating/preventing infection by
 CC bacteria, viruses and fungi, affecting bodily characteristics, e.g.
 CC biorhythms, fertility or metabolism, affecting behavioural
 CC characteristics, and for providing analgesic effects. A host cell
 CC containing the polynucleotide is useful to produce non-human transgenic
 CC animals.
 XX
 SQ Sequence 1851 AA;

Query Match 89.6%; Score 2882; DB 22; Length 1851;
 Best Local Similarity 81.8%; Pred. No. 1.7e-279;
 Matches 581; Conservative 7; Mismatches 32; Indels 90; Gaps 5;

QY 1 MLAQPORLLFLDGADELPAALGPEAAPCTDPPEAASGARVLGGLLSKALLPTALLVTT 60
 DB 267 MLAQPORLLFLDGADELPAALGPEAAPCTDPPEAASGARVLGGLLSKALLPTALLVTT 326
 QY 61 RAAAPGRLQRLSPQCAEVRGSDKDKKKYFKFFRDERRAERAYRFVKEVETLFCF 120
 DB 327 RAAAPGRLQRLSPQCAEVRGSDKDKKKYFKFFRDERRAERAYRFVKEVETLFCF 386
 QY 121 VPFCVIVCTVLRQQLGRLGRLSRTSKTTTSVLLFTSVLSAPVADGPRLOGDLRNL 180
 DB 387 VPFCVIVCTVLRQQLGRLGRLSRTSKTTTSVLLFTSVLSAPVADGPRLOGDLRNL 446
 QY 181 RLAREGVLRRAQFAEKELEQLRSGKVQTLFLSKKELPGVLETEVTVQFIDQSFQFL 240
 DB 447 RLAREGVLRRAQFAEKELEQLRSGKVQTLFLSKKELPGVLETEVTVQFIDQSFQFL 506
 QY 241 AALSYLEDDGVPRTAAGVGTLLRGAQPHSHLVLTTRFLGLLSAERMDIERHFGCM 300
 DB 507 AALSYLEDDGVPRTAAGVGTLLRGAQPHSHLVLTTRFLGLLSAERMDIERHFGCM 566
 QY 301 VSERVQKELRWYOGOGCGVAPVETGAKGLEDTEPEEEEGEENPYPLELLYCLY 360
 DB 567 VSERVQKELRWYOGOGCGVAPVETGAKGLEDTEPEEEEGEENPYPLELLYCLY 626
 QY 361 ETQEDAFVQALCRFPPELALQVRFCRMDVAVLSYCVRCPCAGQALRLISCRVAAQEK 420
 DB 627 ETQEDAFVQALCRFPPELALQVRFCRMDVAVLSYCVRCPCAGQALRLISCRVAAQEK 686
 QY 421 KKSGLGRLQSLGSSSOCT----- 440
 DB 687 KKSGLGRLQSLGSSWLTQLAEPVFRPPCCDICPTPPPPRLLQKAFARVPLNIAP 746
 QY 441 TKOLPASLL-----HPLFQAMTDPLCHLSLTLGSHCKLPDAVCRDLSEA 484
 DB 747 IOPPLRGSLASVERVNVTLAGACGPDPKTHMTDPCHLSLTLGSHCKLPDAVCRDLSEA 806
 QY 485 LRAAPALTELGLLHNLRLSEAGLSEGLAWPOCRVQTVRVQPPQGRLOLVGMJLQ 544
 DB 807 LRAAPALTELGLLHNLRLSEAGLSEGLAWPOCRVQTVRVQPPQGRLOLVGMJLQ 866
 QY 545 PALTTLDLSGQCPAPMWTVLZAVLHOGCGCTLSLA-----SVELSE----- 589
 DB 867 PALTTLDLSGQCPAPMWTVLZAVLHOGCGCTLSLSPSPTSSSGRCRPPGRRLG 926
 QY 589 -----CSLQELQAVKRAKPDLVIT-----HPALCHCHPOPPKE 620
 DB 927 LESRWPRSAPEPSGSEARTQVZAAGGAGGRRRGREPPARGPHQPPFD 976

RESULT 5
 AAO17860
 ID AAO17860 standard; Protein; 1851 AA.
 AC AAO17860;
 XX 20-AUG-2002 (first entry)
 DE Pyrin domain containing protein NALP6/PY9.
 XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
 KW nephroprotic; osteopathic; nootropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis.
 XX OS Unidentified.
 XX PN WO200240668-A2.
 XX 23-MAY-2002. **d**
 XX 30-OCT-2001; 2001WO-EPI2545. **d**
 XX 15-NOV-2000; 2000DE-1056687.
 XX 30-NOV-2000; 2000DE-1059595.
 XX (APOT-) APOTECH RES & DEV LTD.
 XX Tschopp J, Mattinor F;
 XX WPI; 2002-427093/45.
 XX N-PSDB; AAL47132.
 XX New DNA encoding protein with pyrin domain, useful for treating
 XX diseases involving impaired signal transduction, particularly
 XX inflammation, also proteins and antibodies -
 XX Claim 5; Fig 1; 116pp; German.
 XX The present invention relates the DNA and their encoded proteins, where
 XX the proteins contain at least one PYD (pyrin) domain. These can be used
 XX to treat diseases associated with impaired intracellular signal
 XX transduction, particularly inflammation such as psoriasis,
 XX arteriosclerosis, bacterial or viral infections (particularly meningitis
 XX and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 XX sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 XX and Parkinson's diseases. The present sequence is a protein of the
 XX invention.
 SQ Sequence 1851 AA;

Query Match 89.6%; Score 2882; DB 23; Length 1851;
 Best Local Similarity 81.8%; Pred. No. 1.7e-279;
 Matches 581; Conservative 7; Mismatches 32; Indels 90; Gaps 5;

QY 1 MLAQPORLLFLDGADELPAALGPEAAPCTDPPEAASGARVLGGLLSKALLPTALLVTT 60
 DB 267 MLAQPORLLFLDGADELPAALGPEAAPCTDPPEAASGARVLGGLLSKALLPTALLVTT 326
 QY 61 RAAAPGRLQRLSPQCAEVRGSDKDKKKYFKFFRDERRAERAYRFVKEVETLFCF 120
 DB 327 RAAAPGRLQRLSPQCAEVRGSDKDKKKYFKFFRDERRAERAYRFVKEVETLFCF 386
 QY 121 VPFCVIVCTVLRQQLGRLGRLSRTSKTTTSVLLFTSVLSAPVADGPRLOGDLRNL 180
 DB 387 VPFCVIVCTVLRQQLGRLGRLSRTSKTTTSVLLFTSVLSAPVADGPRLOGDLRNL 446
 QY 181 RLAREGVLRRAQFAEKELEQLRSGKVQTLFLSKKELPGVLETEVTVQFIDQSFQFL 240

447 KLAPEGVIGRRRAQFAEKLEQLRGSKVQTLFLSKXKLEPGVLETEVYQFIDQSQFEFL 536
 241 AALSYLEDDGVPTAAGGVTLRGDAQPHSHLVLTTRFLGLLSERNRDIERHFGCM 332
 507 AALSYLEDDGVPTAAGGVTLRGDAQPHSHLVLTTRFLGLLSERNRDIERHFGCM 566
 301 VSERVKQEARLWVGQGGPGVAPVTEGAKGLEDTEEPDEEEEGEPNYPLELYCLY 360
 567 VSERVKQEARLWVGQGGPGVAPVTEGAKGLEDTEEPDEEEEGEPNYPLELYCLY 626
 361 ETQDAFVRQALCFPELALJRVFRCDMDVAVLSYVRCPCPAGQALSLISCRVAAQEKX 420
 627 ETQDAFVRQALCFPELALJRVFRCDMDVAVLSYVRCPCPAGQALSLISCRVAAQEKX 686
 421 KKSIGKRLQASLGSGSSQGT-----SVELSE----- 440
 687 KKSIGKRLQASLGSGSGWLTGLAPEVFRPPCCDICTPPDPRLLGCKAFARVPLNIAP 746
 441 TKQCPASILL-----HPLFQAMTDPLCHLSLTLSHCKLPDAVCRODLSA 484
 747 ICPPLRGLASVERMNVTVLAGPGDPKTHAMTDPLCHLSLTLSHCKLPDAVCRODLSA 806
 485 LRAAPALTELGLLHNRLESEARMLSEGLAWPCQRCVQTVRVQDPDQGGQYLVGMRLQOS 544
 807 LRAAPALTELGLLHNRLESEARMLSEGLAWPCQRCVQTVRVQDPDQGGQYLVGMRLQOS 866
 545 PALTTDLGSCQLPAPMVTYLCVLOHOGGLOTLSLA-----SVELSE----- 588
 867 PALTTDLGSCQLPAPMVTYLCVLOHOGGLOTLSLA-----SVELSE----- 926
 589 -----OSLOFQAVKRAKPOLVIT-----HPALDGHPPQPKPE 620
 927 LESRWPSAPFSGDSEARTJVEAAGGAGGRRRGREPPARGPFPQPRD 976
 RESULT 6
 ABB11735
 ID ABB11735 standard; peptide; 597 AA.
 XX AC ABB11735;
 XX DT i--JAN-2002 (first entry)
 XX DE Human vasopressin receptor homologue, SEQ ID NO:2105.
 XX KW Human: cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW anasthetic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cyostatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;
 KW antifungal; vulnerary; antiulcer.
 XX OS Homo sapiens.
 XX PN WO200157188-A2.
 XX PD 09-AUG-2001.
 XX PF 05-FEB-2001; 2001WO-US03800.
 XX FR 03-FEB-2000; 2000US-0496914.
 XX PR 27-APR-2000; 2000US-0560875.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;

XX KPI: 2001-457740/49.
 DR N-PSDB; ABA08979.
 XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX Claim 20; Page 236; 1963pp; English.
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides,
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, and thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis; cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of virus.
 CC Bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.

Sequence 597 AA;

Query Match: 69.6%; Score 2239; DB 22; Length 597;
 Best Local Similarity 100.0%; Pred. No. 1.3e-215;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIAQPORLLFLIDGADLPALGPPAAPCTDPFEAASGARVLGGLLSKALLPTALLVTT 60
 DB 163 MIAQPORLLFLIDGADLPALGPPAAPCTDPFEAASGARVLGGLLSKALLPTALLVTT 222
 QY 61 RAAAPRQLGRLCSQCAEVRGFSKDKKKYKFFRERRAERAYRVKNETIFALCF 120
 DB 223 RAAAPRQLGRLCSQCAEVRGFSKDKKKYKFFRERRAERAYRVKNETIFALCF 282
 QY 121 VPFVCMIVCTVLRQOLELGRDLRSKTTTSVYLLFITSVLSSAPVADGPRLOGDLRLC 180
 DB 283 VPFVCMIVCTVLRQOLELGRDLRSKTTTSVYLLFITSVLSSAPVADGPRLOGDLRLC 342
 QY 181 RLAREGVLRRAQFAEKELEQLERGSKVQTLFLSKKELPGVLETEVYQFIDQSQFEFL 240
 DB 343 RLAREGVLRRAQFAEKELEQLERGSKVQTLFLSKKELPGVLETEVYQFIDQSQFEFL 402
 QY 241 AALSYLEDDGVPTAAGGVTLRGDAQPHSHLVLTTRFLGLLSERNRDIERHFGCM 300
 DB 403 AALSYLEDDGVPTAAGGVTLRGDAQPHSHLVLTTRFLGLLSERNRDIERHFGCM 462

QY 301 VSERVKQEARVVGQGGCGVAVETGAKGLEDTPEPEEEGEGEPNPJELLYCLY 360
 DB 463 VSERVKQEARVVGQGGCGVAVETGAKGLEDTPEPEEEGEGEPNPJELLYCLY 522
 QY 361 ETQEDAFVQAQCRPELALQVRFCRMDVAVLSYCVRCPCPAGQALRLISCRLVAAQEKK 420
 DB 523 ETQEDAFVQAQCRPELALQVRFCRMDVAVLSYCVRCPCPAGQALRLISCRLVAAQEKK 582
 QY 421 KKSGLKRLQASLGGG 435
 DB 583 KKSGLKRLQASLGGG 597

RESULT 7

AAR33389

ID AAR33389 standard; Protein; 481 AA.

XX AAR33389;

AC AAR33389;

DT 25-MAR-2003 (updated)

DT 15-JUL-1993 (first entry)

XX AII/AVPv2 receptor.

DE AII/AVPv2 receptor.

XX Recombinant; angiotensin II; cAMP; vasopressin-v2; AII; AVPv2; vector;

KW receptor; G protein-coupled; neurotransmitter; angiotensin II; human;

KW arginine vasopressin; isoreceptor; mammal; rat; extracellular domain;

KW immunogen; antibody; AII/AVPv2.

XX Synthetic.

OS Synthetic.

XX Key

FH Location/Qualifiers

FT Peptide

FT Peptide

FT Peptide

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

CC receptors respond to arginine vasopressin. Isoreceptors have been
 CC described for both AVP and AII, this is consistent with their
 CC functional diversity. This recombinant receptor may be derived from
 CC a mammal, preferably a human or a rat. The DNA encoding this sequence
 CC may be used in the construction of vectors for the expression of the
 CC recombinant receptor. Fragments of the extracellular domain of the
 CC receptor are useful as immunogens for producing antibodies which
 CC neutralise the activity of the AII/AVPv2 receptor in vivo.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 481 AA;

SQ Sequence 481 AA;

Query Match

Best local similarity 49.0%; Score 1578; DB 14; Length 481;

Matches 328; Conservative 49; Mismatches 98; Indels 14; Gaps 5;

Cyt 136 LEIGRDLRSRTSTTSVYLLFITSVLSSAPVADGPRQLGDLRLCLAREGLVGRRAQFA 195

DB 1 MELGRDLRSRTSTTSVYLLFITSVLSSAPVADGPRQLGDLRLCLAREGLVGRRAQFA 195

QY 196 EKELEOLELRGSKVOTLFLSKKELPGVLETVTYQF:DQSFQEFLLAALSYLLEDGGVPRPT 255

DB 60 EXDLERLKGSGQVQTMFLSKKELPGVLETVTYQF:DQSFQEFLLAALSYLLEDGGVPRPT 255

QY 256 AAGGVCTLLRGDAQPHSHLVLTTRFLGLLSAERMRDIERHFGCMVSRVXQEAJRWVOG 315

DB 20 SAGSVQMLLSADAGLRGHALTRFLGLLSAERMRDIERHFGCMVSRVXQEAJRWVOG 315

QY 316 QGGCGPGVA---PEVTEGAKGLEDTPEPEEEGEGEPNPJELLYCLYETQEDAFVQAQ 372

DB 180 QSO--PKVATVGAERKDELKDEAEDEEEDEEEDEEEDEEEDEEEDEEEDEEEDEEEDE 237

QY 373 CRPELALQVRFCRMDVAVLSYCVRCPCPAGQALRLISCRLVAAQOE--KKKSLGKSLQA 430

DB 238 SSLEPMVLERVRLTRMDLEVLVSYCVRCPCPAGQALRLISCRLVAAQOE--KKKSLGKSLQA 430

QY 431 SLGGSSQGTTKQLPASLLHPLFQAMTDPDLCHLSLTLSSHCKLPDAVCRDLSEALRAAPA 490

DB 297 -----QSQTGKQPPASLLRPLCEAMITQCGGLSILTLSSHCKLPDAVCRDLSEALRAAPA 490

QY 491 LTELGLLHRLSEAGLRMLSEGLAWPQCVQVVRVQLPDPFQGLQVLMQLRQSPALTTL 550

DB 352 LRELGLLQNLTEAGLRLLSQGLAWPQCVQVVRVQLPDPFQGLQVLMQLRQSPALTTL 550

QY 551 DLSCGCLPAPMWTYLLCAVLCHQCGGLTSLASVLSLSEQLQELQAVKAKPDLVTHPA 610

DB 412 DLSCGCLPAPMWTYLLCAVLCHQCGGLTSLASVLSLSEQLQELQAVKAKPDLVTHPA 610

QY 611 LDGHPQPPK 619

DB 472 LGTHPQPLK 480

RESULT 8

ABB90173

ID ABB90173 standard; Protein; 321 AA.

XX ABB90173;

AC ABB90173;

DT 24-MAY-2002 (first entry)

DE Human polypeptide SEQ ID NO 2549.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

KW antileukemic; hepatotropic; antidiabetic; antiinflammatory; antitumor;

KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW candiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

OS Homo sapiens.

XX WO2002:90304-A2.

PN

XX


```
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239315.
PR 13-OCT-2000; 2000US-0239317.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0250390.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.

PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUXAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-4B3426/52.
XX N-PSDB; AAK62741.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX Claim 11; SEQ ID NO 17553; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX Sequence 329 AA;

Query Match 32.8%; Score 1056.5; DB 22; Length 329;
Best Local Similarity 83.9%; Pred. No. 41e-97;
Matches 219; Conservative 5; Mismatches 34; Indels 3; Gaps 3;

Qy 1 MLAQPORLLFILGADLPALGPEAAPCTDPEAASGARVLGGLSKALLPTALLVTT 60
Dy 66 MLAQPORLLFILGADLPALGPEAAPCTDPEAASGARVLGGLSKALLPTALLVTT 125
Qy 61 RAAAPORLQRLCSPQCAEVGRFSDKKKKYFYKFFRDERRAERAYRFVKNETLPALCF 120
Dy 126 RAAAPORLQRLCSPQCAEVGRFSDKKKKYFYKFFRDERRAERAYRFVKNETLPALCF 185
Qy 121 VPFVCMIVCTVLRQQLGRLDLSRTSKTTTSVYLLFITSVLSSAPVADGPRQGLDRLNLC 180
Dy 186 VPFVCMIVCTVLRQQLGRLDLSRTSKTTTSVYLLFITSVLSSAPVADGPRQGLDRLNLC 245
Qy 181 RLAREGVLGRRQAFAEKELEQLERLRSK-VQTLFLSKKELPGVLETEVTYQFDGQSFQF 239
Dy 246 RLAREGVLGRRQAFAEKELEQLERLRSKADACFSAKRCACWQSRSPSSSTKSFQGV 305
Qy 240 -LAALSYLE-EDGGVPRTAG 258
Dy 306 PRGHLSTLAGRTGGVAGGTAG 326

RESULT 10
AAO15590
ID AAO15590 standard; Protein; 1061 AA.
XX AAO15590;
XX AAO15590;
XX 31-OCT-2002 (first entry)
XX Human PYPIN-8 protein #2.
XX
```

KW Human; gene therapy; PYRIN; stress-related response; apoptotic response;
 KW inflammatory response; inflammatory disorder; immune system disorder;
 KW Crohn's disease; multiple sclerosis; cancer; leukaemia;
 KW autoimmune disorder; arthritis; neurological disease;
 KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
 KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
 KW transcription profiling; PYRIN-8.
 XX Homo sapiens.
 OS WO20026:049-A2.
 PN 08-AUG-2002.
 XX 31-JAN-2002; 2002WO-US02967.
 XX 31-JAN-2001; 2001US-265231P.
 PR 10-SEP-2001; 2001US-318645P.
 XX (MIL-1) MILLENNium PHARM INC.
 PA (AMHP) WYETH.
 PA Bertin J, Wang W, Blatcher M;
 P: MPI; 2002-627477/67.
 DR N-PSDB; AAL44363.
 XX New PYRIN polypeptides and nucleic acids useful for modulating and
 PT diagnosing stress-related, apoptotic and inflammatory responses, or for
 PT treating inflammatory and immune system disorders, cancers, or
 PT neurological diseases
 XX Claim 8; Fig 8; 16pp; English.
 XX The invention comprises the amino acid and coding sequences of human
 CC PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
 CC useful for modulating and diagnosing stress-related, apoptotic and
 CC inflammatory responses. The PYRIN protein and DNA sequences are useful
 CC for treating inflammatory disorders and immune system disorders (e.g.
 CC Crohn's disease, reactive arthritis, multiple sclerosis, contact
 CC dermatitis, psoriasis, graft rejection, allergies, viral infections and
 CC bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
 CC (e.g. systemic lupus erythematosus and arthritis); and neurological
 CC diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
 CC protein and DNA sequences may also be used in screening assays, detection
 CC assays (e.g. chromosomal mapping, tissue typing or forensic biology),
 CC predictive medicine (e.g. diagnostic assays, clinical trials and
 CC pharmacogenomics) and transcription profiling. The present amino acid
 CC sequence represents a human PYRIN-8 protein.
 XX SQ Sequence 1061 AA;
 Query Match 25.7%; Score 825.5; DB 23; Length 1061;
 Best Local Similarity 32.6%; Pred. No. 4.3e-73;
 Matches 219; Conservative 104; Mismatches 245; Indels 103; Gaps 15;
 QY 1 MLAQPORLLFILOGDEL-PALGGPERACTDPERAASGARVLGGLLSKALLPTLLCAVT 59
 Db LIRVPERLLFILOGDELKSFHDHPQSPWLCWEKRPTELLDNLNLRKKLLPELSLT 341
 QY 60 TRAAAPGRLOGLCSQCAE/YRGFSXDKKKYFYKFFRDERRAERAYRFVKENETLFLAC 119
 Db TRPTALEKHLHLEHPRHVEILGFSEAEKRYFYKYFHNAEQGVFNVDNEPLFTMC 402
 QY 120 FVPFVCHIVCTVLRQLELGRDLSRTKTTTSVYLLEITSVLSSAPVAQPRLO--GDLR 177
 Db FVPLVCHVCTVLRQLELGRDLSRTKTTTSVYLLEITSVLSSAPVAQPRLO--GDLR 177
 QY 178 NLCLAREGLVGRPAQFAEKELELRGSKVOTLFLSKXELPCVLETEVYQF--DQSFQ 237
 Db GLCSLAADGLWNLKILFEEQDLRKHGUDGEDV-SAFUNMIFQKINCERYYSFIHLSQ 518
 QY 238 EFLAALSYLEDDGGVPRTAAGVG-----TLRGDAQPHSHLVLTTRFLFLGLLSAERM 290

Db EPPAAVYILDEG-----EGGAGPDODVTRLLTEYAFSERSFLATSRFLFGILLNESTR 572
 QY 291 RDIERHFCMVSRVKQEARLWVGQGGQCPGVAPEVTEGAKGLEDETEPEEEEEGEPPN 350
 Db SHLEKSLCWKVSPIKMDLLQWISKAQ-----SDGSTLQQGS--- 610
 QY 351 YPLELLYCLYETQSDAFVROALCFPELALORVFCRMDVAVLSVCVRCPCPAGQALRLIS 410
 Db --LFFSCLYEIQEEFIQALSHFQVIVSNTA-SKMEHVSVSFCILKRCCKSAQVLEHYG 667
 QY 411 CRLVAAOEKKKSLGKRLQASLGGGS-----SQGTTKQLPASJ----- 448
 Db ATYSADGE-----DRARCSAGATLLVQLPERTVLLDAYSEHLAAALCTNPMLIELSL 720
 QY 449 -----LHPLFOAMTDPPLCHSLSLTSLSHCKLPDAVCRDLSEALRAAPALTSLGSLHN 499
 Db YRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKILMTMDLSGN 780
 QY 500 RLSEAGLRMLSEGLAWPQCRVQTVRVQLPDPQRG-LQVLVGLKLRQSPALTTLDLSGGCLP 558
 Db GVGFPGMLLCEGLRHPCRLQMIQLKRCQLESACQEMASVLGTNPHLVELDTGNAL 840
 QY 559 APVVTYLCVILQHQCGGLQTL-----SLASVELSEQSLQELQAVKRAKPDLV 605
 Db DLGRLLLCQGLRHPVCR-LTLWLKICRLTAACDELASTLSVNSQLRELDLS.NELGDLG 900
 QY 606 I-----THP 609
 Db VLLCEGLRHP 911
 RESULT 11
 AAO17857
 ID AAO17857 standard; Protein; 1099 AA.
 AC AAO17857;
 DT 20-AUG-2002 (first entry)
 XX Pyrin domain containing protein NALP3/PY5-hs.
 DE Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
 KW nephroprotic; osseoprotic; nootropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis.
 XX OS Unidentified.
 XX WO200240668-A2.
 PN 23-MAY-2002.
 PD 30-OCT-2001; 2001WO-EPI2545.
 PF 15-NOV-2000; 2000DE-1056687.
 PR 30-NOV-2000; 2000DE-1059595.
 XX (APCT-) APOTECH RES & DEV LTD.
 XX Tschoopp J, Martinon F;
 DR WP2; 2002-427093/45.
 DR N-PSDB; AAL47129.
 XX New DNA encoding protein with pyrin domain, useful for treating
 PT diseases involving impaired signal transduction, particularly
 PT inflammation, also proteins and antibodies
 XX Claim 5; Fig 1; 116pp; German.

FT Modified-site 595..597
 FT /label= Prote.n_kinase_C_phosphorylation_site
 FT Domain 606..607
 FT /label= Dileucine_motif
 FT Region 618..626
 FT /note= "Perox.somal targeting signal"
 FT Modified-site 624..627
 FT /label= Case.n_kinase_II_phosphorylation_site
 FT Modified-site 654..657
 FT /label= N_glycosylation_site
 FT Modified-site 656..658
 FT /label= Protein_kinase_C_phosphorylation_site
 FT Modified-site 657..660
 FT /label= Casein_kinase_II_phosphorylation_site
 FT Modified-site 717..722
 FT /label= N_myristoylation_site
 FT Modified-site 740..743
 FT /label= Casein_kinase_II_phosphorylation_site
 FT Region 740..767
 FT /label= Leucine_rich_repeat
 FT Domain 740..991
 FT /label= LRR_domain
 FT Modified-site 750..753
 FT /label= Casein_kinase_II_phosphorylation_site
 FT Region 769..796
 FT /label= Leucine_rich_repeat
 FT Region 797..821
 FT /label= Leucine_rich_repeat
 FT Domain 815..816
 FT /label= Dileucine_motif
 FT Region 816..837
 FT /note= "Leucine zipper pattern"
 FT Domain 823..824
 FT /label= Dileucine_motif
 FT Region 826..849
 FT /label= Leucine_rich_repeat
 FT Region 854..878
 FT /label= Leucine_rich_repeat
 FT Modified-site 888..893
 FT /label= N_myristoylation_site
 FT Region 883..906
 FT /label= Leucine_rich_repeat
 FT Modified-site 911..914
 FT /label= N_glycosylation_site
 FT Modified-site 919..924
 FT /label= N_myristoylation_site
 FT Region 911..935
 FT /label= Leucine_rich_repeat
 FT Modified-site 921..924
 FT /label= Casein_kinase_II_phosphorylation_site
 FT Domain 929..930
 FT /label= Dileucine_motif
 FT Domain 934..935
 FT /label= Dileucine_motif
 FT Region 940..967
 FT /label= Leucine_rich_repeat
 FT Modified-site 950..953
 FT /label= N_glycosylation_site
 FT Domain 962..963
 FT /label= Dileucine_motif
 FT Modified-site 968..970
 FT /label= Protein_kinase_C_phosphorylation_site
 FT Region 968..991
 FT /label= Leucine_rich_repeat
 FT Modified-site 970..973
 FT /note= CAMP and cGMP dependent protein kinase phosphorylation site
 FT Domain 997..998
 FT /label= Dileucine_motif
 FT Modified-site 1014..1017
 FT /label= Casein_kinase_II_phosphorylation_site
 FT Modified-site 1018..1021
 FT /label= N_myristoylation_site

XX WO20016:005-A2.
 FX
 XX 23-AUG-2001.
 FX
 PF 16-FEB-2001; 2001WO-US40145.
 XX
 XX 17-FEB-2000; 2000US-0506067.
 PR 01-SEP-2000; 2000US-0653901.
 PR
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 XX Bertin J;
 XX
 XX WPI; 2001-514773/56.
 DR N-PSDB; AAD14323.
 DR
 XX Novel polypeptide comprising NBS-1 or PYRIN-1 of the pyrin domain
 PT protein family useful in screening and detection assays and for
 PT treating, e.g., cancer, viral infections, autoimmune disease, and
 PT Alzheimer's -
 XX
 PS Claim 8; Fig 4; 111pp; English.
 XX
 CC The invention relates to human NBS-1 (nucleotide binding site) and
 CC PYRIN-1 protein and nucleic acids. NBS-1 and PYRIN-1 DNA, protein and
 CC their modulators are useful in the treatment of apoptotic and
 CC inflammatory disorders, cancer (leukaemia, melanoma, carcinoma); viral
 CC infections (including herpesvirus and adenovirus), autoimmune diseases
 CC (systemic lupus erythematosus (SLE), immune-mediated glomerulonephritis,
 CC arthritis); neurological disorders (Alzheimer's disease, Parkinson's
 CC disease, amyotrophic lateral sclerosis (ALS), retinitis pigmentosa,
 CC Huntington's disease); aplastic anaemia, ischaemia, meningitis, liver
 CC diseases. NBS-1 and PYRIN-1 DNA, protein and their modulators are also
 CC used for the treatment of inflammatory and immune disorders such as
 CC chronic inflammatory diseases such as Crohn's disease, insulin-dependent
 CC diabetes, organ specific autoimmunity, including multiple sclerosis,
 CC Grave's disease, allergy, asthma, HIV, tuberculosis and lepromatous
 CC leprosy. The present sequence is human PYRIN-1 protein.
 XX
 SQ Sequence 1034 AA;
 Query Match 24.9%; Score 800.5; DB 22; Length 1034;
 Best Local Similarity 28.4%; Pred. No. 1.3e-70;
 Matches 219; Conservative 117; Mismatches 243; Indels 193; Gaps 17;
 QY 1 MIAOPORLLFLDGADELPAALGGPEAAP-CTDPPEAASGARVGLGSLKALLPTALLVLT 59
 DB 288 IVRKPSRILFLMDGFDLQGAFFDEHIGPLCTDWCKAERGDLISLIRKLLPEASLLIT 347
 QY 60 TEAAPGRLOGLRCSPOCAEVRGSDKKYVYKFFPDERRAERAYFVKENETLFAIC 119
 DB 348 TRPVALEKLOHLDHPRHVEILGFSEAKRKEYFFRYFSEACAAAFSLIGENEVLFTMC 407
 QY 120 FVPFVCWIVCTVLRQOELGRDLRSRTSTTSVLLFTITSVSSAPVADGPRLOCDLRL 179
 DB 408 FIPLVCWIVCTCLKQOMESGSLACTSKTTAVYVFLSSLLQPRGSGQEHGLCAHNL 467
 QY 180 CLAREGVLAGRAQPAEKELEQLRGSKVOTFLFSKKELPGVLETVTYTFIDGSPQEF 239
 DB 468 CSLAADGIWNGKILFEESDLRNHGQKADV-SAFRLMNLFOKEVDECKEYFIFHMTFOEF 526
 QY 240 LAAISLYLLEDGGVPRTAAG-----VGTLLRGDAQ-PHSHLVLTTRFLGSLSER 289
 DB 527 FAAMYVLJEEKEGRTNVPGRSLKLPSSRDVTVLLYENYKFEKGYLIFVVRFLGLVNOER 586
 QY 290 MRDIERHFQCMVSRVQKQALRWYQGGQCGPGVAPEYTEGAKGLEDETEPEEESEGE 349
 DB 587 TSYLEKLSKISQIRLELLKWI-----EVKAKKQLQ--IQSQ----- 625
 QY 350 NYPLELLYCLYETQDAFVQALCRFPPELALQVRFCRMQVAVLSYCVRCPCPAGQALRLI 409
 DB 626 ---LELFYCLYEMQBEDFVQAMDYFPKIE-N--LSTRMDHVMSSFCIENCHRVESLSLG 680

QY 410 SCRLVAAQEKKKSLGKRLQ-----ASLGGSSQG-TTKQ-PASLLHPLF-----CA 455
 Db 681 FJHNPKEEBEKEGRHLDWQCVLPSSSHAACSHGLVNSHLTSSFCRGLFVSLSQS 740
 QY 456 MT-----DP----- 459
 Db 741 LTFELDLSDNSLGDPMGRVLCETLQHPGCIIRRLWLGRCGLSHECCFCISJVLSSNQKAVE 800
 QY 460 -----LHLSLTLSHCKLPDAVCRDLSALRAAPALTGL 496
 Db 801 LBLSDNALGDFGIRLLCVG-KHLLCNLKKLWLVSCCLTSACQDLASVLSHSLRYV 860
 QY 497 LHNRSSEAGRLMSEGLAPQCRVOTYV----- 525
 Db 861 GENALGDSVAALCEKAKPQCNQKLGJVSNGLSVCCSA:SSVLSNQNTHLYLRGN 920
 QY 526 QLPDPORGQYLV-GMLRSPALTTLDLSGCCLPAPM----- 561
 Db 921 T-LGD--KGIKLCEGLLHPCKLQVLELONCNLTSHCCWDLSTLTSSOSLRKLSLGNNE 978
 QY 562 -----VTYLCVAVLOHOGCGLTLSLASVELSQSLQELQAVKXKAPDLVI 636
 Db 979 LGDLGVNMFCEVLKQSGCLQNLGUSEYFYETKSALETLOEERPELIV 1028

RESULT 13
 AB08503
 ID AB08503 standard; Protein; 1034 AA.
 XX AC AB08503;
 XX 22-MAY-2003 (first entry)
 DE Huma PYRIN-1 protein.

XX Human; nucleotide binding site; pyrin domain: NBS-1; PYRIN-1;
 KW caspase-1; antiinflammatory; apoptosis; ASC; NF-kB; nuclear factor KB;
 KW LRR; leucine rich repeat; inflammatory disorder; familial cold urticaria;
 KW arthritis; inflammatory bowel disease; Crohn's disease; asthma; allergy;
 KW ulcerative colitis; rheumatoid arthritis; Lyme disease; Grave's disease;
 KW insulin-dependent diabetes; multiple sclerosis; contact dermatitis;
 KW psoriasis; graft versus host disease; food allergy; conjunctivitis;
 KW chronic obstructive pulmonary disease; viral infection; HIV infection;
 KW helminthic infection; Leishmaniasis; bacterial infection; tuberculosis; leprosy; chromosome 1q44.
 XX Homo sapiens.
 OS
 XX US2002187922-A1.
 XX 12-DEC-2002.
 XX 22-APR-2002; 2002US-0.27516.
 XX 17-FEB-2000; 2000US-0506367.
 XX 31-SEP-2000; 2000US-0653901.
 XX 26-SEP-2001; 2001US-0964955.
 XX 20-DEC-2001; 2001US-0027629.
 XX (BERT/) BERTIN J.
 XX (MANU/) MANU J G A.
 XX Bertin J, Manji GA;
 XX MF: 2003-328763/31.
 XX N-PSDB; ABX93556.
 XX Identification of compound that binds to polypeptide (for e.g. a PYRIN
 XX protein), useful for treating disorders associated with inappropriate
 XX apoptosis, for e.g. inflammatory disorder -
 XX Claim 1; Fig 4; 74pp; English.

XX The invention relates to a new method for the identification of a
 CC compound that binds to a human PYRIN-1 (an apoptotic signalling
 CC molecule appearing as AB08503) comprising contacting the polypeptide or
 CC a cell expressing the polypeptide to a test compound and determining
 CC whether PYRIN-1 binds to the test compound, or determining the effect of
 CC the test compound on the activity of PYRIN-1. In a similar manner,
 CC compounds are isolated which modulate the binding of PYRIN-1 to ASC
 CC (comprising testing the compounds against the binding of PYRIN-1 to ASC
 CC both PYRIN-1 and ASC (not defined)), modulate the activity of NF-KB
 CC (nuclear factor KB), and which modulate the ASC-mediated activation of
 CC NF-KB (comprising: (a) measuring the binding of a test compound to the
 CC LRR (leucine rich repeat) domain of PYRIN-1; and (b) measuring the
 CC activation of NF-KB in a cell expressing ASC and PYRIN-1 in the presence
 CC and absence of the LRR domain binding compound). The candidate
 CC modulators may be useful for treating an inflammatory disorder.
 CC Also included is a method (M8) for identifying a modulator of caspase-1
 CC activity. The identified compound can be used to modulate the
 CC polypeptide's activity (ASC and NF-kB activities in a patient). By
 CC modulating the expression or activity of the polypeptide (PYRIN-1 and
 CC ASC), a disorder associated with inappropriate apoptosis (inflammatory
 CC disorders e.g. familial cold urticaria, arthritis, inflammatory bowel
 CC disease, Crohn's disease, ulcerative colitis, rheumatoid arthritis, Lyme
 CC disease, insulin-dependent diabetes, multiple sclerosis, Grave's
 CC disease, contact dermatitis, psoriasis, graft versus host disease,
 CC asthma, chronic obstructive pulmonary disease, allergies (e.g. food
 CC allergies), conjunctivitis, helminthic infection (e.g. Leishmaniasis),
 CC viral infections such as HIV infection, and bacterial infections such as
 CC tuberculosis and leprosy) can be treated. Human NBS-1 protein (nucleotide
 CC binding site protein 1) contains a pyrin domain which was used to
 CC identify genomic sequences containing the PYRIN-1 gene. The gene for
 CC PYRIN-1 is located on chromosome 1q44. The present sequence represents
 CC human PYRIN-1.
 XX
 XX Sequence 1034 AA;
 SQ

Query Match 24.9%; Score 800.5; DB 24; Length 1034;
 Best Local Similarity 28.4%; Pred. No. 1.3e-70;
 Matches 219; Conservative 117; Mismatches 241; Indels 193; Gaps 17;
 QY 1 MLAQPORLIFLDGADLPALGGPEAP-CTDPFAASGARVLGGLSKALLPTALLVT 59
 Db 288 IVKPSRIPLFMDGDFELQGFDEHIGLCTQWKAERGDIILLSSLRKKLPESALLIT 347
 QY 60 TRAAAPGRLQGLRCSPOCAEVGFSKDKKFKYFPFRDERAERAYRVRKNETLFLALC 119
 Db 348 TRPVALEKQLHLLDHRHVEILGFSEAKKEFFKVFSEDAQARAASLIQNEVLFTMC 407
 QY 120 FVPFVCMIVCTVLRQQLGRDLRSKTTTSVYLLFTSVLSAPVADGPRLOGLRL 179
 Db 408 FIPLVCMIVCTGLKQOMESGKSLAQTSKTTTAVYVFLSSLLQPRGSGQEHGLCAHLMEL 467
 QY 180 CELAREGVLGRRAQFAEKELEQLERLGSKVQTLFLSKKELPGVLETEVYQIFDOSFOEF 239
 Db 468 CSLAAGIWNQKILFEESDLRNHGLQADV-SAFLLMNLFOKEVCEKFKYSIHMTFQEF 526
 QY 240 LAALSYLLEDGGVPRTAAG-----VGTLLRGDAQ-PHSHLVLTTRFLGLLSAER 269
 Db 527 FAAMYLLSEKRGRTNVFGRKLPSRDVTVLENYGKFGYLIFFVVRFLGLVQNER 586
 QY 290 MRDIERHFGCMVSRVKQALRWVQGGCGCPVAPETVTEGAKGLEDETEPEEEGEP 349
 Db 587 TSYLERKLSCKISQIRLELLKWI-----EVKAKAKLQ--IQPSQ----- 625
 QY 350 NYPLELLYCLYETOEDAFVROALCRFPPELALQVRFCRMDVAVLVSVRCPPAGALRLI 409
 Db 626 ---LEJFYCLYENQESDFVQRANDYFPKIEIN--LSTRMDHMYSSFCIENCHRVESLSLG 680
 QY 410 SCRLVAAQEKKKSLGKRLQ-----ASLGGSSQG-TTKOLPASLLHPLF-----QA 455
 Db 681 FLHNPKEEBEKEGRHLDWQCVLPSSSHAACSHGLVNSHLTSSFCRGLFVSLSQS 740
 QY 456 MT-----DP----- 459

```

Db 741 LTELSDNSLGPGRVRLCTLQHPCCNIRLWLGRGLSHCECCDTSVLSSNOKLVE 800
QY 460 -----LCHSSLTLSHCKLPDAVCRDLSEALRAAPATFELGJ 496
Db 801 LLSLSDNALGFGIRLLCVGJGHLLCNKKLWLVSQCC:TSACCODLASVLSHSLRLVY 860
QY 497 LHNRLSEAGLRLMSEGLAWPQCRVQTVRV----- 525
Db 861 GENALGDSGVAILECEKAKIPQCQLKGLVNSGLTSVCCSALSSVLS:TNQNTHLVLRN 920
QY 526 QLPDPGRGLQYLV-GWLRQSPAT-TLGLSGCQ:PAFM----- 561
Db 921 TCGD--KGILCEGLJHPDCKLVLEEDMCN:TSHCWDLSTLTSSOSLKLSLGNND 978
QY 562 -----VTYLCVAVLHQCGGLQTLTSLASVELSEOSLQELQAVKRAKPDVLI 606
Db 979 LGDLGVMWFCEVLKQOSCLLQN:LGLSEMYENYETKSALET:LQEEKPELV 1028

RESULT i4
ID AAB62571
AC AAB62571;
XX 23-JUL-2001 (first entry)
XX Human CARD-7 polypeptide.
DE CARD-7: CARD-5; caspase recruitment domain; cancer; human;
KW autoimmune disorder; anti-inflammatory; immunosuppressive; antiallergic;
KW antibacterial; antiviral; gene therapy.
XX Homo sapiens.
XX Key Location/Qualifiers
PH Domain 329..645
FT /note= "nucleotide binding domain"
FT Domain 333..341
FT /note= "kinase 1A subdomain"
FT Domain 404..413
FT /note= "kinase 2 subdomain"
FT Domain 454..463
FT /note= "kinase 3a subdomain"
FT Domain 615..622
FT /note= "motif 2 domain"
FT Domain 807..814
FT /note= "leucine-rich domain"
FT Domain 836..863
FT /note= "leucine-rich domain"
FT Domain 864..891
FT /note= "leucine-rich domain"
FT Domain 893..920
FT /note= "leucine-rich domain"
FT Domain 921..948
FT /note= "leucine-rich domain"
FT Domain 950..976
FT /note= "leucine-rich domain"
XX WO2001308:3-A1.
XX 03-MAY-2001.
XX 27-OCT-2000; 2000WO-US29796.
XX 27-OCT-1999; 99US-0428252.
XX (MILL-) MILLENNIUM PHARM INC.
XX Bertin J;
XX WPI; 2001-343340/36.

```

```

DR N-PSDB; AAF83651.
XX identifying a modulator of interaction between caspase recruitment
PT domain (CARD)-7 and CARD-5, for treating autoimmune disorders,
PT comprises measuring the binding of CARD-7 and CARD-5 in the presence of
PT the compound.
XX Disclosure; Fig 1A-D; 80pp; English.
XX The invention relates to identifying compounds, that modulate interaction
CC between caspase recruitment domain (CARD)-7 and CARD-5. The method
CC involves measuring the binding of CARD-7 and CARD-5 in the presence of
CC the compound (an increase in the binding of CARD-7 to CARD-5 in the
CC presence of the compound compared to the binding in the absence of the
CC compound indicates that the compound is a modulator of CARD-7-CARD-5
CC interaction). Modulators of CARD-7 and CARD-8 expression or activity can
CC be used to treat or diagnose disorders such as cancers, bacterial or
CC viral infections, autoimmune disorders (systemic lupus erythematosus,
CC immune-mediated glomerulonephritis or arthritis), inflammatory disorders,
CC organ-specific autoimmunity including multiple sclerosis, Hashimoto's
CC thyroiditis, or Grave's disease, psoriasis, graft rejection, allergies.
CC CARD-7 and CARD-8 are useful as modulating agents in regulating a variety
CC of cellular processes including cell growth and cell death. The present
CC sequence represents the human CARD-7, an intracellular protein.
XX Sequence 1429 AA;
SQ Query Match 23.7%; Score 762; DB 22; Length 1429;
Best Local Similarity 33.2%; Pred. No. 1.6e-66;
Matches 220; Conservative 103; Mismatches 207; Indels 132; Gaps 22;
QY 1 MLAGPQRLFLFDGADLPP--ALGGPEAAPCTDPEAASGARVLGGLSKALLTALLLV 58
DB 398 LSRPERLLFLFDGVE-PGWVLQEPSELCHWSQFPADALGSLGKLTILPEASFL 456
QY 59 TTRAAAPGRLOGLCSPCAEVRGFSQDKKKYKFFROERRAERYFVKENETPAL 118
DB 457 TARTALONLIPSEQARWVEVLGFSESSKEYRYFTDERQAIRAFLVKSNKELWAL 516
QY 113 CFVPFVCMIVCTVLRQQLLEGRDLSRTSKTTTSYVLLFITSVLSSAPVADGPRJQGBLRN 178
DB 517 CLVPWVSWLACTCLMQQMKRKEKLTLSKTTTLLHYLAQALQAQPL--GPQ---JRD 570
QY 179 LCRLAREGVLRRAQPAEKELESLQELSGSKVQTLFLSKKELPGVLETE---VTQFIDQS 235
DB 571 LCSLAAGINQKTLFSPDRLRGHLDGAI:ST-FLKY---GILQHPILPSYSFIHLC 625
QY 236 FOEFLAALSYLEDDGVFRTAAGGVGVTLLRGDAQPHSHLVL-----T 277
DB 626 FOEFFAAMSIVJED-----EKRG-----KHSNCIIDLKTLKAYGIHGLFGAS 672
QY 278 TRFLFGLLSAERMRIERHFQCMYSERVKQELRWVQGGCGPGVAPVETEGAKGLED 337
DB 671 TRFLGLLSDEGEREMENIFHCRLSQ--GRNLMQWV-----PSLQLL 710
QY 338 EEPPEEEGEEPNYPLEL:YCLYETQEDAFVROALCRFPPELALQVRFC---RMDVAVL 394
DB 711 LQP-----HSLESGLCHLYETRNKTLFTQVMAHFEEMGM-----CVETDELLVCT 755
QY 395 YCVRCPCPAGQALRLISCR-----LVAAGEKKKSLGKRLQ 429
DB 756 FCIKFSRHVKLQLLEGQRHRTWSP:WVVLFRWVPVTDAYWQILFSLVKVTRN:KELD 814
QY 430 ASLGGGSGQTTKOLPASLLHPLFOAMTDPCHLSSSLTSLSHCKLPDAVCRDLSBALRAAP 489
DB 815 LS---GNS-----LSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAECDCKDLAFGLRANQ 865
QY 490 ALTELGLLHNLSEAGLRLMSEGLAWPQCRVQTVRVQLPQPRGL-----CYLVGM:RQS 544
DB 866 TLTELDLSFNVLTDAGAKHLCCQLRQPSCKLQ--RLQL--VSCGLTSDCCODLASVLSAS 921
QY 545 PALTTLDLGGCQLPAPMYTVYLCVAVLHQCGGLQTLTSLASVEL:SEQSLEQLQAVKRAKPD 604

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 17:45:33 ; Search time 107 Seconds
(without alignments)
1507.317 Million cell updates/sec

Title: US-10-029-347-2

Perfect score: 3218
Sequence: 1 MIAQORLLFLDGADFLPA.....ITHPALDGHPPPKELISTF 625

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:
1: sp_arChea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_protist:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	835	25.9	1033	11 Q8R4B8	Q8R4B8 mus musculus
2	829.5	25.8	892	4 Q8NEU4	Q8NEU4 homo sapien
3	644.5	20.0	713	6 Q95LZ7	Q95LZ7 macaca fasc
4	569.5	17.7	846	4 Q8XT0	Q8XT0 homo sapien
5	473	14.7	982	11 Q8BU40	Q8BU40 mus musculus
6	464.5	14.4	748	11 Q99NM0	Q99NM0 mus musculus
7	429.5	13.3	825	11 Q8C6M5	Q8C6M5 mus musculus
8	412.5	12.8	863	11 Q8C6J9	Q8C6J9 mus musculus
9	369.5	11.5	657	11 Q8EPG7	Q8EPG7 mus musculus
10	356	11.1	673	11 Q8GCN1	Q8GCN1 mus musculus
11	321.5	10.0	692	4 Q96J51	Q96J51 homo sapien
12	305	9.5	1020	11 Q8K3Z0	Q8K3Z0 mus musculus
13	289	9.0	953	4 Q5W5F5	Q5W5F5 homo sapien
14	271.5	8.4	778	4 Q8NF48	Q8NF48 homo sapien
15	267.5	8.3	1155	7 Q9TPP1	Q9TPP1 mus musculus
16	254.5	7.9	706	11 Q8BUT6	Q8BUT6 mus musculus

17	254.5	7.9	953	11 Q8BH80	Q8BH80 mus musculus
18	238	7.4	1052	7 Q9GJD8	Q9GJD8 rattus norv
19	238	7.4	1073	7 Q9GJD9	Q9GJD9 rattus norv
20	238	7.4	1153	7 Q9GJE0	Q9GJE0 rattus norv
21	235.5	7.3	519	11 Q8C249	Q8C249 mus musculus
22	235	7.3	660	4 Q8NF06	Q8NF06 homo sapien
23	233	7.2	884	7 Q23675	Q23675 homo sapien
24	233	7.2	932	4 Q96KL4	Q96KL4 homo sapien
25	229	7.1	509	4 Q9H5Z8	Q9H5Z8 homo sapien
26	226	7.0	1056	4 Q8NF42	Q8NF42 homo sapien
27	223	6.9	977	7 Q8HW99	Q8HW99 mus musculus
28	213	6.6	461	6 Q8HZP9	Q8HZP9 pan troglod
29	212	6.6	447	4 Q96FD7	Q96FD7 homo sapien
30	212	6.6	461	4 Q8ZKX8	Q8ZKX8 homo sapien
31	211	6.6	461	4 Q8BQ80	Q8BQ80 homo sapien
32	211	6.6	516	4 Q8NBF5	Q8NBF5 homo sapien
33	210.5	6.5	287	4 Q9BY26	Q9BY26 homo sapien
34	206	6.4	1097	4 Q9H6Y0	Q9H6Y0 homo sapien
35	198.5	6.2	456	11 Q91VI7	Q91VI7 mus musculus
36	196.5	6.1	456	11 Q924P4	Q924P4 mus musculus
37	18C	5.6	390	11 Q86WZ1	Q86WZ1 mus musculus
38	170	5.3	1004	5 Q944Z7	Q944Z7 cichla intes
39	169.5	5.3	733	4 Q8TEE2	Q8TEE2 homo sapien
40	162	5.0	554	4 Q9H724	Q9H724 homo sapien
41	149	4.6	753	4 Q969L7	Q969L7 homo sapien
42	146	4.5	269	11 Q9D458	Q9D458 mus musculus
43	134	4.2	1194	4 Q9H7K4	Q9H7K4 homo sapien
44	130	4.0	237	11 Q8BT74	Q8BT74 mus musculus
45	130	4.0	743	13 Q9JGM4	Q9JGM4 gallus gall

ALIGNMENTS

RESULT 1

ID Q8R4B8 PRELIMINARY; PRT; 1033 AA.
AC Q8R4B8; 21, Created)
CT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Mast cell maturation inducible protein 1.
GN MMIG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/cJ;
RA Kikuchi-Yanosita R., Koga K., Taketomi Y., Sugiki T., Saito T.,
RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;
RT "Identification of inducible genes during in vitro maturation of mouse
RL bone marrow-derived mast cells to connective tissue-type mast cells.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF486632; AAL90874.1; -;
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR007091; LRR_RNinh.
DR InterPro: IPR003590; LRR_RNinh_sub.
DR InterPro: IPR007111; NACHT_NTPase.
DR InterPro: IPR004020; PAAD_DAPIN_dom.
DR Pfam: PF00560; LRR; 2.
DR SMART: SM00368; LRR_R1; 1.
DR PROSITE: PS00824; DAPIN; 1.
DR PROSITE: PS00036; HLH_1; 1.
DR PROSITE: PS00503; LRR_R1; 3.
DR PROSITE: PS00837; NACHT; 1.
SQ SEQUENCE 1033 AA; 118274 MW; 5924690966B1217 CRC64;
Query March 25.9%; Score 835; DB 11; Length 1033;

Best Local Similarity 29.3%; Pred.No.le=57;
Matches 226; Conservative 112; Mismatches 239; Inserts 194; Gaps 15;

```

QY      1  MLAGQFRLFIIDGACGELP-ALGGPEAPACTDPPEAASGARVLGGILSKALLPTALLVNT 59
Db      1  :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
286  ILAKFSRILFMDGFDGQAGAFDEHIGEVCTDWQKAVRGDILLSTLRKDLLPKASLUT 345
QY      60  TRAAPGRLOGRLCSPQCAEVRGSDKKKFKYKFFRDEPRAPRAYRFFKENTFTFLALC 119
Db      60  :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
346  TRVALEKLOHLLDHPRHVELTGFSEAKKEYFFKYFSNELQAEAEARLQENEVLTNYC 405
QY      120  FVPFVCHIVCTVLRQQLEGRDLRSTSKTTSVLLFITSVLSSAFVADQPRLOGDENL 179
Db      120  :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
406  FIPLVCHIVCTGLKQOMETQKSLAQTSTTAVYVFFLSSLLQSRGGTTEHLFSCYLQGL 465
QY      180  CRLAREGVLCRRAPFAEKELQLELRGSKVOTFLSKKELFGVLETEVTOFIDQSQOEF 239
Db      466  CSUADGIRKQKILFECDDIRKGGKQKTDV-SAFIRXNVPOKEVDCRRFVSFKHXTQEF 524
QY      240  LAALSYLE---DGVVPRTAAGVGITLLRGDAQ-----PHSHVLVLTTRFFGLLSA 287
Db      525  FANMYV:LEEEAEGETVRKGGGCSLLNRDVKVLRLENYKFKEXGLIFVVRFLFGLVNG 584
QY      288  ERNRDIERHFGCMVSRVKQOELRWVGQGGCGVAPETVTEGAKGLEDEETPEEESEGE 347
Db      585  ERTSY:LEKLSCKISQOVRLELLKWI-----EVKAKAKKLQ--WQPSQ----- 625
QY      348  EPNVPLELLVCLYETQEDAFVROALCFPELALQVRFCRDVAVLVSVRCRCPAGQALR 407
Db      626  -----LEFLVCLYEMQBEDFVQSAMDHFFKIEIN--LSTRYDHWVSSFCIKNCHRVKTL 678
QY      408  L:SCRLVAAQEKKKKSJGKRLQ-----ASLGGG--SSQGTTK 442
Db      679  LGFHNSPKEEEERRGRPLDQVQCVFPCTHVACSSRLVNCCLTSSFCRLGFSLSLNR 738
QY      443  QJ-----PASLHPLFQAMTDPCHLSSLTSLSHCKLPDVCRODLSALRAAPA 490
Db      739  SLTELDLSDNTLGDPG--MRVLCEALQHPGCNIQRLWLGRCGLSHQCCFDISSVSSSQK 796
QY      491  LTELGLLHLNRLSEAGRLMSEGLAWPOCR----- 519
Db      797  LVELDLSDNALGDFGIRLLCGVJGKHLKCNLQKMWVCCU:TSACCCQDALVLSNHSLETR 856
QY      520  -----VQTVRVQLPDCRGQYLL----- 537
Db      857  LYIGENALGSGVQVLCERKMKPCNQLKGLVNSGLTS:CCSALTSVLKTNQNFTLYL 916
QY      538  -----VGMVROSPTALTIDJSGCQLPAP----- 560
Db      917  RSNALGDTGIRLCEGLLHPDCKLQMLDELNDCSLTSWSCNWLSTILTHNHSRLKRLNGNN 976
QY      561  -----VYTYLCVAVLQHCQGG:QTLTSLASVELSEQLCELGAVKRAKPDQVI 606
Db      977  DLGDLCVTLCEVLKQGGCLQSLQGLXGYLNRETKEALEALQAEKEPELTI 1027

RESULT 2
Q8NEU4
ID Q8NEU4 PRELIMINARY; PRT: 892 AA.
AC Q8NEU4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DD 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Monarch-1 splice form IV.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_Taxid=9606;
RN [1]
RP
RA Williams K.L., Linhoff M.W., Ting J.P.Y.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY116207; AAW75145.1;

```

DR	InterPro; IPR007C9.; LRR_RNinh.
DR	InterPro; IPR007C11; NACHT_NTPase.
DR	InterPro; IPR004C20; PAAD_DAPIN_dom.
DR	Pfam; PF02758; PAAD_DAPIN_1.
DR	PROSITE; PS00824; DAPIN; 1.
DR	PROSITE; PSS0503; LRR_RI; 1.
DR	PROSITE; PSS0637; NACHT; 1.
SQ	SEQUENCE 892 AA; 101733 MW; AE703D8DF341C2AC CRC64;

Query Match	25.88;	Score 829.5;	DB 4;	Length 892;
Best Local Similarity	33.18;	Pred. No. 2.3e-57;		
Matches 215;	Conservative 109;	Mismatches 241;	Indels 85;	Gaps 14;

Cy	1	MLAQPCRLLLFDGADSL-PALGPEAAPCTDPPEAASGARVJLGLLSKALLPTALLLVIT	59
Dd	282	LIRVPERELLFIIDGFDELKSFHDPPQCWCLWEKEPTELLNSLRKLLPELSLIIT	341
Cy	60	TAAAPGRLOGRLCSPCAVRFGSDKKKKYFYKFYFRERRERAERAFVKENETFLFALC	119
Dd	342	TRPTALEKLHLLLEHRPHVEILGFSAEKEYEYKYYPHNAEQAGOVFNVDNEPLFTMC	401
Cy	120	FVLGVCHIVCTVLROOLEGRDLRSRTSKTTTSYLLFITSVLSARPAVDGRQL--GDLR	177
Dd	402	FVLVCHVVCTCLQQQLGGGLLRQTSTTAVVMYLXLLSMQPKPGA--PRLOPPPNNR	459
Cy	178	NLCRLAREGYLGRRAQFAEKELGLELRGSKVOTFLSKKELPGVLETENVYQIFDSFQ	237
Dd	460	GLCSLAADGLWNQKILEEGDLRKHGJDGEDV-SAFLNMAIFDKINCERYYSFHLISFQ	518
Cy	238	EFLAASLYLEDGGVPRTAAGVG-----TLRGDAQPHSHLVLTTRFVGLLSAERM	290
Dd	519	EFFAAMYIIDEG-----EGGAGPDQVTRLLEYAFSERSFALTISRFLUGLLNESTR	572
Cy	291	RDIERHFGCMVSERVKOEALRWVQGQCGGCPVAPEVTGNKG--EDTEEPEEEBEGEPN	350
Dd	573	SHLEKSICWKVYSPHKMDLLOWIQSKAQ-----SDGSTLQOGS---	610
Cy	351	YPLELLYCYETOEDAFVROALCRFPPELLQRVRFRCRMVDVAVLVCYVRCPCPAGALRLIS	410
Dd	611	--LEFFSCLEIQDEEFIQALSHFQVIIVSNIA-SOMEHWVSFCCLKRCSAVLHYLG	667
Cy	411	CRLVAAOEEKKXSLGRLEQASLGGS-----SOOTTKOLPASL-----	448
Dd	668	ATYSADGE-----DRARCAGAHTLLVOLRPRTVLLDAYSEHLAAALCTNPNIELS	720
Cy	449	-----LHPLFQANTDPLCHLSLTIHQKLPDAVCEDLSLARAAAPALTGLGLH	498
Dd	721	LYRNALSGRVKCLCOGLRHPNCKLQNLKRKRRISSACEDSLAAIANKNKTRMCDLSG	780
Cy	499	NRLSEAGRLMSLEGAWPQCRVOTVRVOLPDPORG-LQYLVGMREQSPALCTLDLSGCQL	557
Dd	781	NNGVFPYMWLLCEGLRHPQCRQLQWLQRLKCKLESAGCAQEMASVLGTNPHLVELDITGNAL	840
Cy	558	PAPMTVYLCVAVLQHGGGLQTLTSLASVELSEQSLQELQAVKRAKPDIVI	606
Dd	841	EDLGLRLCCGLRHPVCRLTRLMWLFQMGLNKMTHSRLAALRVTKPYLDI	890

RESULT 3				
Q95LZ7	PRELIMINARY;	PR7;	713 AA.	
AC Q95LZ7	01-DEC-2001 (T-EMBrel. 19, Created)			
DT 01-DEC-2001 (T-EMBrel. 19, Last sequence update)				
DE 01-MAR-2003 (T-EMBrel. 23, Last annotation update)				
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;				
OC Cercopithecinae; Macaca.				
OX NCBI_TaxID=9541;				
RP SEQUENCE FROM N.A.				
RP [1]				

IC	Q8N
AC	Q8N
DT	01-
DT	01-
DT	01-
DE	Mon
OS	Homo
OC	Euk
OC	Mamm
OX	NCBB
RN	[1]
RP	SEQ
RA	Will
RL	Sub
DR	EMB


```

RESULT 7
Q8C6M5 PRELIMINARY; PRT; 825 AA.
AC Q8C6M5;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Weekly similar to PAN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK054264; BAC3570.1; -.
KW Hypothetical protein.
SQ SEQUENCE 825 AA; 94021 MW; 11D71DEABEAFCA31 CRC64;

Query Match 33.3%; Score 429.5; DB 11; Length 825;
Best Local Similarity 25.4%; Pred. No. 1.7e-25; Indels 135; Gaps 21;
Matches 163; Conservative 105; Mismatches 238;

QY 6 QRLFLDGADEL--PALGGPZAAFTCD-----PFEAASGARVLGGLLSKALLPTALLV 58
DB 254 KEFLFVWDGDELTPA-GALIRLDGDMWTKVPEV-----LLGSLKRWKAPHAALLV 307
QY 59 TTRAAAPGLOGRLSPQCAEVRGSDKDKKYYFKFRDER-----RAGRAYRV 109
DB 308 TTRTOSLHQIFVMDQPLIVETLGFLEQEKQYFKYFEDEEGEEDGEGKARLAKVEV 367
QY 110 KENETLFCFVPVCTVLRQOELGRDLRSKTTSTTVYLFTITSVLSAPVAD- 168
DB 368 RGNADLYQWASPTAGGFCLELRMMKGGELSLTCQYTSFMFLNFCVFSSETCEDH 427
QY 169 -GPRIQGLDRLNCRILAREGVLRRAQFAEKELELRGSKVQTL-----FLSKKELPGVL 223
DB 428 LNEEFQILPKK-CILAANSLLRQVILCEDEPFLTKLNLNHPMVCRHILFK----- 480
QY 224 EFTVY--QFIDQSQEEAALSYLEEDG----GYPRTAAGVGTLIRGDAQ-PHSKVL 276
DB 481 DSSSTHCLSFICLGIQQLLAAIFVQELQESKGVSKYS---IQNMJSREARLNKPNPLSG 537
QY 277 TTRFLFGLSABRMRIERHFGCMVSERVKQBALRWVQGGCGCPGVAPEVTEGAKGLED 336
DB 538 LLPFVFGLLNETRIQELKTTFCQISTEVKRFKFL----- 571
QY 337 TEEPEEEEGEENPVL-----ELLVCLYTQDQAFVQALCRFPPELALQVRFCRMD 399
DB 572 -----ECESGE--NKPLLLMMQBIISCLYSOEGFVKAMVLFEDISLHLK--TSTD 622
QY 390 VAVLSYVRCPCPAGQALRLISCRVAAOEKKKSLGKRLQASLGGSSQGTTKQLPASLL 449
DB 623 LIHASEFLK--NSQLQTMSLKVEKA----- 646
QY 450 HPLFQAVTTPCLCHLSLTLSHCKLPD-----AVCRDLSEALRAAPALTELGLJHNRUSEA 504
DB 647 -----VFPENVALES-TAKQRSPEDEQMLTFWTFDCTFNSKKLVF.LDIHESFLNS 700
QY 505 GLRMLSEGLAWPQCRVCTVRVQLPQORGLQVYLVGMLRQSPALTTLDLGGCQLPAPMVTY 564
DB 701 ALLEICEKLPSASCCCKQVVLNKHSPDDAYEKLCLIPNGYKTIISHLILQGGNLDLS-MHHS 759
QY 565 LC AVLHOHCGGLQTLASVELSEOSLOELQAVKZAKPDV 605
DB 760 LCEVLKNPACNLKFLSLGSCSTAAQKWDFFPVLKVNOSLI 80C

```

```

RESULT 8
Q8C6J9 PRELIMINARY; PRT; 863 AA.
AC Q8C6J9;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Hypothetical RNI-like structure containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK054426; BAC35775.1; -.
KW Hypothetical protein.
SQ SEQUENCE 863 AA; 99955 MW; D3A46C782DA544CF CRC64;

Query Match 12.8%; Score 412.5; DB 11; Length 863;
Best Local Similarity 22.6%; Pred. No. 4e-24; Indels 221; Gaps 19;
Matches 165; Conservative 100; Mismatches 245;

QY 1 MLAQPORLLFDGADLP-ALGGPEAAFPCTDPFEAASGARVLGGLLSKALLPTALLV 59
DB 213 ILSQEKLFIDSLGKEMWNTQQDSQCLYCNMCKQPNVLLSLLRKKILPESLLIS 272
QY 60 TRAAPGRLOGLRCPQCAEVRGSDKDKKYYFKFRDERAERAYRVKNETLFAIC 119
DB 273 TSCETFKDLKDMIEVTNVRTITGFKENNINMCFSLFQDRNIAQEAFAFSLIRENEQVFTVC 332
QY 120 FVPPVCMVCTVLRQOELGRDLRSKTTSTTVYLFTITSVLSAPVADGPRLOQD-LRN 178
DB 333 QAVVCVMYATCLKNEIBSGKDPVSCRITTSYTHILNLFIPHNAQPNNSNEDLDN 392
QY 179 LCLRLAREGVLRRAQFAEKELELRGSKVQTLFLSKKELPGVLETEVTFYFIDQSFQ 238
DB 393 LCLFAYEGWMTDISVFNNEEARLNGIMDSIPTL-LDIGILQESRESEKSYFLHPSVQE 451
QY 239 FLAALSYLEEDGVPRTAAGVGTLIRGDAQPHSHL----- 274
DB 452 FCAMFVYLL-----HSEMDHSCQGVYFIETFLFTFLNKKIKKQ 488
QY 275 -VLTTRFLGILLSABRMRIERHFGCMVSERVKQBALRWVQGGCGCPGVAPEVTEGAKG 333
DB 489 WYFLGCFPGLLHETQEKLEAFPGYHLSKELRRQLFLWLE-----L 530
QY 334 LEDTEEPBEEEGEENPYLELLCYLYETQDQAFVQAL-CRFPPELALQVRFCRMDVAV 392
DB 531 LLDTLHPDVKKIN-----TMKFFCYCLFEMEVEEVFQSAMNCRE-----QIDVVV 574
QY 393 LSV-----CVRCPCPAGQALRLISCRVAAOEKKKSLGKRLQASLGGSSQGTTKQLPASLL 427
DB 575 KGYSDFIVAAYCLSHGSAUTDFS---ISAQNVINBELGQRGKLLILMHQICSVFLRNKDI 631
QY 428 -----LQASLGGSSQGTTKQLPASLL----- 449
DB 632 KTLRIEDTIFNPVFKIFYSYLNKSSCILKTAVNVSVFLCDKRLFLELIQSYNLSLEYL 691
QY 450 -----HPLFQAVTTPCLCHLSLTLSHCKLPD-----AVCRDLSEALRAAPALTELGLJH 499
DB 692 RGTFLSHSDVEMLCOTLNQAEQNRILDLANGSLCEHSDYLSVLRQNKSLRYLNISYN 751
QY 500 RLSEAGLRMLSEGLAWPQCRVCTVRVQLPQORGLQVYLVGMLRQSPALTTLDLSSCQLPA 559
DB 752 NLKDEGLKAL-----CRAUT-----IP-----NSALHSUSLEACQITG 784

```

```
QY 560 PM-----VTYLCVAVQHOGGLOTLISLASVELSEQS- 593
Db 785 ACKDLASTFTRKCLRRINLAKNSJGFGSLFVLCMAKQDCTLYELKLRMAFDSDSQ 844
QY 591 ---LQELQAVK 598
Db 845 EFJLSEYERK 855

RESULT 5
Q9EPG7 PRELIMINARY; PRT; 657 AA.
AC Q9EPG7
CT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE RNI-like protein.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10290;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX XEDLINE=21310002; PubMed1416212;
RA Lane R.P., Cutforth T., Young J., Athanasios X., Friedman C.,
RA Rowen L., Evans G., Axel R., Hood L., Trask B.J.;
RT "Genomic analysis of orthologous mouse and human olfactory receptor
RT loci."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7390-7395(2001).
DR EYEL; AF32:233; AAG45188.1;
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR PROSITE; PS50503; LRR_R1; 1.
DR PROSITE; PS50837; NACHT; 1.
SQ SEQUENCE 657 AA; 75410 MW; F4BED9E4BA19AA3 CRC64;

Query Match 11.58; Score 369.5; DB 11; Length 657;
Best Local Similarity 25.8%; Pred. No. 7.4e-21;
Matches 126; Conservative 83; Mismatches 194; Indels 85; Gaps 11;

QY 135 QLELGRDLSRTSKTTTSVLLF-TSVLSSAPVADGPRLOQD--LRMLCLRLAREGVLSRA -92
Db 163 QVEGGRVETCTKTSTALFTYICSLFPVPVG-CVTLNETLRLSLCAVGEIWTXK 221
QY 193 QFAKEKELELELGSKVCTLFLSKKLELPGVLETVYQF-DQSFQBFJAALSYJ-----LE 245
Db 222 VLYQONLRKHELAREDI-LFLDAKVLQCCTEYENCYMF-HLHVQBFPAALFY-LLEENLE 280
QY 249 DGGVPTAAGVGTLRGDAQPHSHAV-TTRELFGLLSARMDIERHSGCMYSEVKCE 306
Db 281 EQYPSPEPFENYJYLDESNT-HDPHLEQMKCFUGLLNKDRVQJQJLNTLITSMXEVRE 340
QY 309 ALRWVQGGGCGCPVAPETVTEGAKLEDTEPEEEEGEEEPNYPLELLYCLYTQEDAFV 368
Db 341 LLALESU-----EKDDSSLSQLRFQDLHLCIYEQDEP 376
QY 369 QALCRFPPELALQVRFCRMJ-----VAVLSYCVRCPCPAGQALRLISCLVAAEKKKSL 424
Db 377 TOALMYFQKIV-----RVJEEPQLRIYSGLGHCHTLXIMLTA-----RADLKNY 423
QY 425 GKRLQASLGGSSGGTTKQLPASLHPLFOAMTDPPLCHLSLTLSHCKLPDAVCRLDSEA 484
Db 424 LDTAEMCLEGAUV-----VH-----VWQDLFSV 448
QY 485 LRAAPALTELGLLHNLRLSEALRLMSEGLAPQCRVQTVRQVLPQFQGLVYLVGLKRS 544
Db 449 LHTNESLJEMDLYESRLCES-MKILNELSHPKCKLOKLIFRSVQFLNGQDFT-FLASN 557
QY 545 PAULTTDLSCQQLPAPMVYTCVAVLQHOGGLOTLISLASVELSEQSLEQELQAVKRAKPDJ 604
Db 508 KKVTHLDLQKTDLGWGLKLT-CEALKCKGCKLRVLRLASCDNLNVARCQKLSNALQTNRSJ 567
```

```
QY 605 VITHPALD 612
Db 568 VFLNLSLN 575

RESULT 10
Q9CCN1 PRELIMINARY; PRT; 673 AA.
AC Q9CCN1
CT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE PAAD and NACHT containing protein..
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK032446; BAC27872.1; -.
SQ SEQUENCE 673 AA; 76367 MW; BE773C592BEC7054 CRC64;

Query Match 11.1%; Score 356; DB 11; Length 673;
Best Local Similarity 25.8%; Pred. No. 9.1e-20;
Matches 118; Conservative 78; Mismatches 174; Indels 88; Gaps 10;

QY 1 MAQPRQLFLIDGADELPAALGPEAAPCTDPFAAGARVLGGLSKALLPTALLVTT 60
Db 233 ILRQPRGLFLIDGVDELQK--SSRAEC-----VLHILMRREVPCS-LJTT 277
QY 61 RAAPAGRLQGLCSPOCAEVGFSDDKKKKYKFFEDERRAEARAYRVKENETLFCALCF 120
Db 278 RPPALQSLPEMLGERRHVLVLGFSSEERETYSFSCFTDKQKNALEFQKNAVLKACQ 337
QY 121 VPFVCWIVCTVLRQOELGRDLSRTSKTTTSVLLFITSVSSAPVADGPRLOQD--LRN 178
Db 338 VPGICWVCSWLKXKXARGQEVSTPSNSTDIFAYVSTFLPTDNGSDSELTRHKVLKS 397
QY 179 LCLRLAREGVLRRAOFAEKLEQLLELGRSKVOTLFLSKKLELPGVLETVYCFIDQSFQ 238
Db 398 LCSLAAEGMRHQRLLFEFEVLRKHGLDGPSTL-TAFNLNCIDYRAGLIGIKKYSFRHISFO 456
QY 239 FJAALSYLLEDGGVPRTAAGVGTLRGDAQPHSHLVTTLFGLLSAERYROIREFEG 298
Db 457 FFYAMSFVLVKDQSQGGQGEATHEKAVLVDPENHEVTLSQLQFDMLKTESGLSLGLKF- 515
QY 299 CMVSRVQKEALRWVQGGCGCPVAPETVTEGAKLEDTEPEEEEGEEEPNYPLELLY 358
Db 516 -----CFRIAPSVR-----QDLKHKEQIEAIKYKRSWGLEFS 548
QY 359 LYETOEDAFVQALCRFPPELALQVRFCRMVAVLSYCVRCPCPAGQALRLISCLVAAQ 419
Db 549 LYDSK-----IKKLTQGIQMKDVLNVQCHJDE 575
QY 419 K---KKKSLGKRLQASLGG-----GSSQGTTKQLPAS 447
Db 576 KKSDDKKKSVSTTSFSSGKVSQSPFLGNDKSTRKOKKAS 613

RESULT 11
Q96D51 PRELIMINARY; PRT; 692 AA.
AC Q96D51
CT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
```


DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update);
 DE Similar to hypothetical protein FLJ21478.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN 1) SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone;
 RA Strausberg R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BCC3199; AAH13199.2; ..
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh_sub.
 DR SMART; SM00368; LRR_R1; 6.
 DR PROSITE; PS0503; LRR_R1; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 692 AA; 76206 MW; AD9D10E7F4A7BDDE CRC64;

 Query Match 10.0%; Score 321.5; DB 4; Length 692;
 Best Local Similarity 24.0%; Pred. No. 5.3e-17;
 Matches 174; Conservative 85; Mismatches 255; Indels 211; Gaps 27;

 QY 1 MLAPQORLLFDGADLP---ALGGPEAACTDPPEAASGARVZGGLSKALLPTALL 57
 DB 1 MAAAGSHLLFVHGTEHLNLCFRLAG--TGLCSDPPEPOEPAIIVNLRLKVMCPQASIL 58
 QY 58 VTTAAAPGRLGRLCSPOCAEVGSDKDKKFKYKFRDE-----PRAE 103
 DB 59 VTTTPSAIGRIPSKYVG-RYGEICGSDNLQKLYFQLRLNPYCGYVGGSGVATPAQ 117
 QY 104 RAY-----RFVKNETLFCALFVPCVWCIVCTVLRQQLGRLDLSRTSKTTSVYLPI 157
 DB 118 RCHLVOMLSRNLEGHQIAACFLPSYCVLCAT----LHF-LHPTPAGOTVTSIYTSFL 173
 QY 158 T-----SVLSSAPVAGGPRLOGDLRLNLCRLAREGVLRRAQFA-----EKE 198
 DB 174 RLNFSGETLDSTDPNSLMAVAARTMGKLAEGVSSRKTYFSEEDVSCLEAGRTSEE 233
 QY 199 LEQELRGSKVQTLFLSKKELQVLSTVTVYFIDGSPQEFLLAAYLLEDDGGVPTAAG 258
 DB 234 FQLHIFRDRARFFLAPCPEPG---RAGTFVFTVPMQEYLAALYIV---GLRKT-LQ 287
 QY 259 GVG---TLRQDAQPHSHLVL-----TTRFLFGLLSAERMED-ERHFGCMVS--- 302
 DB 288 KYGKEVAELVGRVGEDVS-VLGIKAKLLPLRALPL-FNLIKV-----VPRVFGVMVKS 342
 QY 303 ERVKQE-ALKRWQGGCGGCPGVAPEVTEGAKGLE-----DTEPEEEEGE----- 347
 DB 343 EAVACAMVLEMFEEDYNDVVDQMGASILGVGPRRHPDEPPEDEVFELSPFMFGGLL 402
 QY 348 -----EPNYPLELLCYETODEAFV 368
 DB 403 SAHNRVLAQLGCPINKLDALENNAQAIKKLGLKGRQV-PPSELCHLFFHVFQNRQFS 462
 QY 369 RQALCRFPPEALQVRP-----CRMDVAVLSYCVRCPPAQQ-----ALRL-SCRLVAAQEK- 419
 DB 463 AEVLSLSRLQNLNLAGVRMTPVKCTVVAALVG-----SGRHALDEVNLASCQLDPAG-LRT 515
 QY 420 -----KKSLGKRLQASLGSSQSTTQKLPASLLHPLFOAMTDPDLCHLSSTLSHCKL 473
 DB 516 LEPVFLRARKLGLQLN-SLG-----PEACKDLRLDLH-----DOCGITTLRLSNPL 562
 QY 474 PDVACRDLSEALRAAPALTEGLGHLNRLSEAGRLMSEGLAWPQCRVQTVQCLDPDQPG 533
 DB 563 TEAGVAVLMGLAGNYSVTHLSLHTGLDGEGLLELAAGL-----DRNRQ 607
 QY 534 LQYVGMRLRQSPALTTDLSGGQOLPAPMTYLCANV-CHQGGC-LQTLASVAVLSEQLQ 592
 DB 608 LOEL-----KVAYNGAGDTAALARAAREHPSLE 637
 QY 593 ELQAV 597

2b 638 ILQGV 642
 RESULT 12
 Q8K3Z0 PRELIMINARY; PRT; 1020 AA.
 AC Q8K3Z0
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CARD15.
 GN CARD15.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 1) SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Iwanaga Y., Davey M.P., Martin T.M., Planck S.R., DePriest M.L.,
 RA Baugh M.M., Suing C., Rosenbaum J.T.;
 RT "Cloning, sequencing and expression analysis of the murine Nod2/Card15
 gene";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF520774; AAM76073.1; ..
 DR MGI; MGI:2429397; Card15.
 DR InterPro; IPR001315; CARD.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR001130; TcdD_DNase.
 DR PROSITE; PS02029; CARD; 2.
 DR PROSITE; PS0503; LRR_R1; 1.
 DR PROSITE; PS00837; NACHT; 1.
 DR PROSITE; PS01137; TcdD; 1.
 SC SEQUENCE 1020 AA; 113561 MW; 25504905ECF70FEB CRC64;

 Query Match 9.5%; Score 305; DB 1; Length 1020;
 Best Local Similarity 23.5%; Pred. No. 1.8e-15;
 Matches 157; Conservative 111; Mismatches 277; Indels 124; Gaps 28;

 QY 1 MLAPQORLLFDGADLPALGGPEAACTDPPEAASGARVZGGLSKALLPTALLVTT 60
 DB 347 LLDHPQVLLTFDGLDEKFRFTDRERHGS-PIDPTSVQTLFLNLLQGNLLKNAKCVLTS 405
 QY 61 RAAAPGFLGRLCSPOCAEVGSDKDKKFKYKFRDERRABRAYEFVKNETLFCALCF 122
 DB 406 RPAVSALLRKFRTE-LQKGFSEEGIQILYLRKHREPGVADRLLQLQATSAHLGLCH 464
 QY 121 VPFVCMIVCTVLRQQLGRLDLSRTSKTTSVYLLFITSVL-----SSAPVADGPRLL- 172
 DB 465 LPVFSMVVSRCHRELLQNRGPTTS---TDMYLLIQHFLHASPDPSPGLGPGQLLQ 521
 QY 173 --OGDLNLCRLAREGVLRRAQFAKE-LEQLRGSKVQTLFL--SKKELPGVLETVT 228
 DB 522 SRISTLLHGLHLAGRLAMSCTVFSQAQLQAQVSDDISLGLFLVRAQSSVPG---SKAP 578
 QY 229 YQFDIOSFOEFLAA-----LSYLLEDGGV-----PRTAAGG 259
 DB 579 LEFLHITFCFFPAFYLAVSADTSVASLKHFLSCGLSGSLGLLRLPLNLCIQSRVKKGS 638
 QY 260 VTGILLGDAQPHSHLVITTRFLFGLLSAERMRIERHFGCMVSERVYKQPALRWVQGGG 319
 DB 639 EAALLO-KAEPH-NLQITAAFLAGLLS-QOHRDLA--ACQVSRV-----LLQORQARAS 689
 QY 320 CPGVAPETEGAKGLEDTEPEEEEGEPNYPLELLCYETODEAFVRLQALCRFPPELA 379
 DB 690 C--LAHSLREHFHSIPPAVPGETKSMHMPGF-IWLIRSLYEMQEEQLAQAEVRRLL-DIG 745
 QY 380 LQVRVRCRM---DVAVLSYCVRCPPCAQALRLISCRVAAQEKKKSLG----KRLQASL 432
 DB 746 HLKLTFCRCPAECALAFVLQHLQRPVALQL-----DYNVSGDVGVQELRPLCL 794
 QY 433 G-----GGSSOGTTKQLPASLLHPLFOAMTDPLC-HLSSLTLSHCKLPDAVCRDL 481

```
Db 795 GVCTALYLRNNISDRGARTVECAL-----RCECLOKALFNNKULTACACSM 843
Qy 492 SEALAAPALTEGLLHNRLSEAGRLMLSEGLA-----WPCQVGVTVVQVLPDPQ 53:
Db 844 AKLIAHQNFSLRVRGNHITAAGAEVLAQGLKSNSTSLKFLGFWGN-----SVGD-- 893
Qy 532 RGLQYLVGMRLQSPALTTLESGCOLPAPWVTVCAVLOHQGCGLOTLASLVELSEQSL 59:
Db 894 KGTQALAEVADHONLRWLSVGNRI--GSGABALALMLEKNIKSLBELCEENHICDEGV 952
Qy 592 QELCAVKKR 599
Db 953 YSLAEGLEKR 961

RESULT 13
Q81WF5
ID Q81WF5 PRELIMINARY; PRT; 953 AA.
AC Q81WF5
DT 01-MAR-2003 (TremBrel. 23, Created)
DT 01-MAR-2003 (TremBrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBrel. 23, Last annotation update)
DE Caspase recruitment domain family, member 4.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC T-SSUE=Lymph.
RA Strausberg R.;
RL Submitted (NCV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC040339.1; -.
SQ SEQUENCE 953 AA; 107671 MW; 0A9DF167B897E21A CRC64;

Query Match 9.0%; Score 289; DB 4; Length 953;
Best Local Similarity 24.2%; Pred. No. 3.2e-14;
Matches 167; Conservative 91; Mismatches 276; Indels 156; Gaps 27;

Qy 1 MJAQPCRLILDGADLFA---LGGPEAAPCTDPFEAAGARVLGGLSKALLPALLL 57
Db 272 LARFPFHVALTFDGLDELHSLDLSRVPDSSC--PWEPAPHPVLLAKLNLGKLLKGASK 329
Qy 59 VTRRAAPGLOGRLCPQCAEVGFSDDKKYKFFKFFDERRAERAYRFVKNETLFA 117
Db 330 LTAQ---GIEVPQFLRKQYLLRGFSPLHRAVARYFFPERALCDRLSLQAEANPLCS 386
Qy 118 LCFVFPVCMIVCTV---LRQLELGRDLRSRTKTTTSVYLLFITSV-----L 161
Db 387 LCSVPLFCMIIFRCFQHFRAAFEGSPQPDCTMTLTDVFL-VTEVLRNMQPSSLVQRN 445
Qy 162 SSAPVAGPRLQGDRLNLCRLAREGVGRRAQFAEKSEJGLELRGSKVQTLFLSKKELP- 220
Db 446 THSPVETLHAGRTDLCSLGVAHSGMEKSLVFVTQEEVQASGLCDERQYQLGFL--EALPE 503
Qy 221 -GVLETETVTCFIDQSPCEFLAALSYLEO-----GGVPRGAA----- 257
Db 504 LGPGGDDQSEFFHLTLQAFAPFVLDLRVGTQELLRFQEMPPAGAAATSCVPPFL 563
Qy 258 -----GGVGTLLRGDAQFHSVLVLTTRFLGOLLSAERMRIERHFGMVSRVQEAALRW 312
Db 564 PFQCLQSGPAREDLFRNKHOFQTNFLCGLLSKAKQK-LRLHVPAAALRRKRAL-W 622
Qy 313 -----VOGQCGGCPGVAPEVTEGAKGLEDTEPEEESEEPNYPLELLYCYETOED- 365
Db 622 AHLFSSLRGLYKSLPRVQVESFNQVQAM-----PTF-IMMLRCIYETQSQK 666
Qy 366 --AFVRQALCFPELALQVRFCERMDVAVLSYVRCPCPAGQALEL--SCLRLVAQEKKKKS 423
Db 667 VGQLAARGICA-NYLYKXTYCNACSDAGCSALSVLHHFPKRLALDLDNNIL----- 715
Qy 424 LGRQLQASLGGSGSGTTKQLPASLLHPLFQAMTDPCLHLSLTLSHCKLPDAVCRDLSE 483
```

RESULT 14

```
Q8NF48
ID Q8NF48 PRELIMINARY; PRT; 778 AA.
AC Q8NF48
DT 31-OCT-2002 (TremBrel. 22, Created)
DT 31-OCT-2002 (TremBrel. 22, Last sequence update)
DT 31-MAR-2003 (TremBrel. 23, Last annotation update)
DE FJ00348 protein (Fragment).
GN FJ00348.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK090431; BAC03412.1; -.
DR InterPro; IPR001395; Aldo/ket red
DR InterPro; IPR000566; Lipoclin_CytFABP.
DR InterPro; IPR007091; LRR_RNIRh.
DR InterPro; IPR003590; LRR_RNIRh sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR SMART; SM00368; LRR_R1; 2.
DR PROSITE; PS00063; ALDO-KETO-REDUCTASE_3; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS00503; LRR_R1; 1.
DR PROSITE; PS00837; NACHT; 1.
FT NON_TER 1
SQ SEQUENCE 778 AA; 85297 MW; A1EAAQ43FF32836C CRC64;

Query Match 8.4%; Score 271.5; DB 4; Length 778;
Best Local Similarity 24.2%; Pred. No. 6e-13;
Matches 142; Conservative 87; Mismatches 208; Indels 149; Gaps 26;

Qy 2 LAQORLLILDGADLDEL-PALGGPEAAPCTDPFEAAGARVLGGLSKALLPALLLVTT 60
Db 225 VAVPARALLILDGDECRTPLDFTSNTVACTDKKXELPVDHLITNIRGNLPEVSIWITS 284
Qy 61 RAAAPGRLOGRLCSPQCAEVGFSDDKKYKFFKFFRDER-----RAERAYRFVK 11C
Db 285 RPSASGQIPGGGLVD-RMTETIRGFNEEIKVCLQEQFPEDQALLGWLMSQVQADRA----- 338
Qy 111 ENETLFAICFPVFCWIVCTVLRQLELGRDLRSPTS-----KTTTSVYL-FITSVL 161
Db 339 ----LYLMCTVPAPCRLTQMAJG---HLWR--SRTPQDAELMPPRTLCESYWFRLAL 389
Qy 162 S-----SAP-----VADGPRLOGLDLNLCRLAREGVLRRAQFAEKSELE-----QLELR 205
Db 390 SGEQGEKQKASPRISQVAVHGGKRM--VGTGLRLAFHGLIKKKYVEYEQDKAFGVDLALL 447
Qy 206 GSKVOTFLSKKELPGVLETEVTVYQIDQSQEFLAALSY-----LLEDGGV--P 253
Db 448 QCAPSCFLQREE---TLASSVAYCFTHLSLQEFVAAAYVYGASRRRAIFDLFTESGVSWP 504
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: October 2, 2003, 17:46:03 ; Search time 29 Seconds
(without alignments)
911.872 Million cell updates/sec

Title: US-10-029-347-2
Perfect score: 3218
Sequence: 1 MLAQPORLLFLDGADELPA.....ITHPALDGHPPPKELISTF 625

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310f58 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCUS-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	304.5	9.5	1130	2	US-08-519-547A-6
2	289	9.0	953	4	US-09-099-041A-8
3	289	9.0	953	4	US-09-245-281-8
4	289	9.0	953	4	US-09-207-359B-8
5	289	9.0	953	4	US-09-340-620A-8
6	251.5	7.8	953	4	US-09-245-281-43
7	251.5	7.8	953	4	US-09-207-359B-43
8	251.5	7.8	953	4	US-09-340-620A-43
9	224.5	7.3	966	4	US-09-207-359B-47
10	216	6.7	456	2	US-08-910-731-2
11	214	6.7	456	2	US-08-795-395-2
12	214	6.7	456	2	US-08-910-731-8
13	212	6.6	463	2	US-08-910-731-6
14	209.5	6.5	456	2	US-08-910-731-4
15	209.5	6.5	456	2	US-08-795-395-4
16	152	4.7	490	4	US-09-099-041A-26
17	151	4.7	490	4	US-09-245-281-26
18	151	4.7	490	4	US-09-207-359B-26
19	151	4.7	490	4	US-09-340-620A-26
20	137.5	4.3	483	4	US-09-904-615-154
21	126	3.9	200	4	US-09-099-041A-11
22	126	3.9	200	4	US-09-245-281-11
23	126	3.9	200	4	US-09-207-359B-11
24	126	3.9	200	4	US-09-340-620A-11
25	113.5	3.5	759	2	US-08-637-759B-89
26	113.5	3.5	759	3	US-08-871-355A-89
27	113.5	3.5	759	4	US-09-201-945-89

28	104.5	3.2	633	4	US-09-252-991A-21659	Sequence 21659, A
29	104.5	3.2	1129	4	US-09-252-991A-23927	Sequence 23927, A
30	103	3.2	740	3	US-09-022-983-5	Sequence 5, Appl
31	103	3.2	921	4	US-09-252-991A-20327	Sequence 20327, A
32	103	3.2	1411	4	US-09-252-991A-23628	Sequence 23628, A
33	102.5	3.2	617	4	US-09-252-991A-22318	Sequence 22318, A
34	102	3.2	905	4	US-09-340-620A-52	Sequence 52, Appl
35	99	3.1	2509	1	US-08-469-005A-10	Sequence 10, Appl
36	98	3.0	803	4	US-09-154-750A-85	Sequence 85, Appl
37	98	3.0	2511	3	US-09-261-907-2	Sequence 2, Appl
38	97.5	3.0	2756	1	US-08-375-709-11	Sequence 11, Appl
39	97.5	3.0	2756	1	US-08-752-929-11	Sequence 11, Appl
40	97.5	3.0	2756	3	US-09-090-793-7	Sequence 7, Appl
41	97.5	3.0	2756	4	US-09-231-899-7	Sequence 2, Appl
42	97	3.0	912	3	US-08-943-768-2	Sequence 2, Appl
43	97	3.0	912	4	US-09-865-960-2	Sequence 2, Appl
44	96.5	3.0	977	4	US-09-302-812-2	Sequence 2, Appl
45	96.5	3.0	977	4	US-09-511-477-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-58-519-547A-6
; Sequence 6, Application US/08519547A
; Patent No. 5994082

; GENERAL INFORMATION:
; APPLICANT:

; TITLE OF INVENTION: Proteins Essential for the Expression of
; TITLE OF INVENTION: Vertebtrate MHC Class II Genes, DNA Sequences Encoding Same
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10020-1104

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/519,547A
; FILING DATE: 25-AUG-1995
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP94113378.7
; FILING DATE: 26-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, JAMES F.

; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VOS-11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090

; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1130 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-519-547A-6

Query Match 9.5%; Score 304.5; DB 2;
Best Local Similarity 25.5%; Pred. No. 1.9e-23;
Matches 174; Conservative 80; Mismatches 262; Indels 167; Gaps 29;

[illegible]

```

RESULT 2
US-C9-099-041A-8
; Sequence 8, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Hertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TYPE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 67334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-099-041A-8

; Query Match: 9.0%; Score 289; DB 4; Length 953;
; Best Local Similarity 24.2%; Pred. No. 7.1e-22;

```

Matches	167;	Conservative	9;	Mismatches	276;	Indels	156;	Gaps	27;
---------	------	--------------	----	------------	------	--------	------	------	-----

Cy	:	MLAQPORLLFILDGAGELPA--LGGPEAAPCTDPFEAASGARVLGLLSKALLP-ALLI	57
Dd	:	272 LLRFFPHVALFTFDGLDELHSDLDLSRPDSSC--PWEZAPHVLLANLLSGKLLKGAKSL	329
Cy	:	58 VITRAAPQRLQRLCSPOCAEVRGSDKKXKYFYKFPRDERARERAFVVENETLFA	117
Dd	:	330 LTART---GIEVPQRFLRKVKVLLRGFSPLRAYARMFPERALQDRLLSQSEANPNCS	386
Cy	:	118 LCPVPFVMIVCV--LRQQLGRLDLRSCKTTSVVLLFTSV-----L	161
Dd	:	387 LCSVPFLFCIIIFCFPHFAAPEGSQOLDCTNLTLDVFLL-VTEHLNMQPSSLVQRN	445
Cy	:	162 SSAPVADGRLOGDLRNLCRLAREGVGLRRAPAFAEKLEQLERUGSKVOTLFUSKLKP-	220
Dd	:	446 TRSPVELTHAGRTDTCSLQGVAHRGMKEKSLFVPTQBEVOASGLQERJMQLGFL--RALPE	503
Cy	:	221 -GVLETVTYQFIQQSFOEFALALSJJED-----GGVPRTAA-----	255
Dd	:	504 LGPGGDQSVEFHILQAFFTAFFLVDDRVCQELLRFQEWMPAGAA-TSCYPPFL	563
Cy	:	259 -----GGVGTLRGDAQPSHLVLTTRFLGILLSAERMEDIERHFGCMYSERVKQEALRW	312
Dd	:	564 PFQCLQSGSPAREDLFNKDHFQFTNLFLCGLLSKAKQK-LLRHLVPAALREKRKAL-W	621
Cy	:	313 -----VQGGQCPCGVAEVETGAKGLEDETFEEBEGEENYNPLELLYCLYEQED-	365
Dd	:	622 AHUFSSLRGYLSLPVQVESFNQVQM-----PTF-IWLRUCIYE-QSQK	666
Cy	:	366 --AFVFRQALCRFPPELALQRVRCRMVDVLSYCVRCPCAGQALRLISCRVAAQEKKKKS	423
Dd	:	667 VGOLAARGICA-NYLKLYCNACSDCSALS FVLHHFKRELALDDNNL-----	715
Cy	:	424 LGRKLOASLOGGSSQGTTLQLPASLHPHFQAMTDPLCHLSSTLSHSCKLPDAVCDLSE	483
Dd	:	716 -----NDYGVR-----LQPCFSRLT-----VLRLSN-QITDGKVYLSE	750
Cy	:	484 ALRAAPALTGLGLLNRLNSEAGLRMLESEGAWPQCRVQTVRVOLPD----PQRGLQYL---	537
Dd	:	751 ELTKYKIV-YLGIYNNQITDVARYTKIL--DECKGLT-HLKLGKNKITSEGGKYLALA	807
Cy	:	538 -----VGM-----LRQSPALTTLDLSCGQLPAPMYTYLCVLICHQ	572
Dd	:	808 VKNSKSISEYMGNGVGDGAKAFEAALRNHPSLTTLSLASNGISTEGGKS-ARALI-Q	866
Cy	:	573 GCGLOTLTSLASVSEL---CSLQELQAVKR	599
Dd	:	867 NTSLEILWLTONELNDEVAESLAEMLKVNQ	896

RESULT 3
JS-05-245-281-8
: Sequence 8, Application US/09245281:
: Patent No. 6369196
: GENERAL INFORMATION:
: APPLICANT: Berlin, John
: TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
: TITLE OF INVENTION: AND USES THEREOF
: FILE REFERENCE: 07334/118001
: CURRENT APPLICATION NUMBER: US/09/245,281
: CURRENT FILING DATE: 1999-02-05
: EARLIER APPLICATION NUMBER: US 09/207,359
: EARLIER FILING DATE: 1998-12-08
: EARLIER APPLICATION NUMBER: US 09/099,041
: EARLIER FILING DATE: 1998-06-17
: EARLIER APPLICATION NUMBER: JS 09/019,942
: EARLIER FILING DATE: 1998-02-06
: NUMBER OF SEQ ID NOS: 44
: SOFTWARE: FastSeq for Windows Version: 4.0
: SEQ ID NO 8
: LENGTH: 953
: TYPE: PRT

DEC 11 1973

```

RES001 3
US-05-245-281-8
; Sequence 8, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APP-CANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CA
; FILE REFERENCE: 0734/118001
; CURRENT APPLICATION NUMBER: US/09/245,281
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: US 09/207,359
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 09/099,041
; EARLIER FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: JS 09/019,942
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Versior. 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT

```

; ORGANISM: Homo sapiens
US-09-245-291-8

Query Match 9.0%; Score 289; DB 4; Length 953;
Best Local Similarity 24.2%; Pred. No. 7.1e-22;
Matches 167; Conservative 91; Mismatches 276; Indels 156; Gaps 27;

```
QY 1 MLAGPQRLLELDGADLPA---LGGPEAAPCTDPPEAASGARVLGGLLSKALLPTALLL 57
DQ 272 LLRPPHVALFTFDGLDELHDLSDLSRVPDSSC--PWEPAHPLVLLANLLSGKLLKASKL 329
QY 58 VITRAAAPGRLOGLRCSPOCAEVRGSDKDKKFKYKFFRDERRAERAYRVKVENETLPA 117
DQ 330 LTART---GIEVPRQFLRKVKLVLRGSPSHLRAYARMFPERALQDRLLSLEAFNLCS 386
QY 118 LCFVPFVCMIVCTV---LRQLELGRDLRSRTSTTTSVLLFTTSV-----L 161
DQ 387 LCSVPLFCWIIIFRCQHFRAAFEGSPQLPDCPTMTLTDVFL--VTEVHLNMQPSSLVQRN 445
QY 162 SSAPVADGPELQGLRLNCLAREGLVGRRAQPAEKELECLE-RGSKVOTLELSKELP- 220
DQ 446 TRSPVETLHAGRODTLCSLGGVAHGRGKSLFVFTQEVQASGLQERDMQJGFL--RALPE 503
QY 221 -GVLETEVTYQFIDQSFQEFIAALSYLEL-----GGVPRTAA----- 257
DQ 504 LGPGGDCQSYEFHLLTQAFFTAFFLVDDRVTGTOELLRRFFQEWMPAGAAATTCYPPFL 563
QY 258 -----GGVGTLLRGDAQPHSHLVLTTRFLFGLLSAERMRDIERHFGCMYSERVKQEARLW 312
DQ 564 PFQCLQSGGPAREDLFKNDHFQFTNLFLCGLLSKAKQK-LLRHLVPAALRRKRAL-W 621
QY 313 -----VOGQGGQCPGVAPEVTEGAKGLEDTBEPSEEEGEPNYPJELLYCLYETQED- 365
DQ 622 AHLFSSLRGVLKSLPRVQVESFNQVQAM-----PTF-IWMLRCIYETQSQK 666
QY 366 --AFVRAQALCRFPFELALQVRFCRMDVAVLSYVCRCCPAGQALRLISCRILVAQEKKKX 423
DQ 667 VGQLAARGICA-NYLKLTTCNACSADCSALSFVLHHPKRLALCLDNNL----- 715
QY 424 LGKELQASLGGSSQGTQKLPASLLHPLFQAMTDPDLCHLSLTLSSHCKLPDAVCLDSE 483
DQ 716 -----NDYGVR-----LQPCFSRLT-----VLRSLVNQITDGGVKVLSE 750
QY 538 -----VGM-----LRQSPALTTLDLSCGLPAPMVTVLCAVLQHQ 572
DQ 808 VKNSKTSSEVGMGNQVGDGAKAFEAALRNHPSLTTLASNGISTEGGKS-LARALQ-Q 866
QY 573 GCGLOTLASVELSE---OSLOELQAVKR 599
DQ 867 NTSLEILWLTQNELNDEVAESLAEMLKVNQ 896
```

RESULT 4

US-09-207-359B-8

; Sequence 8, Application US/09207359B

; Patent No. 6469140

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED

; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF

; FILE REFERENCE: 07334-112001

; CURRENT APPLICATION NUMBER: US/09/207,359B

; PRIOR FILING DATE: 1998-12-08

; PRIOR APPLICATION NUMBER: US 09/399,041

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: US 09/319,942

; PRIOR FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 953

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-207-359B-8

Query Match

Best Local Similarity 24.2%; Pred. No. 7.1e-22;

Matches 167; Conservative 91; Mismatches 276; Indels 156; Gaps 27;

```
QY 1 MLAGPQRLLELDGADLPA---LGGPEAAPCTDPPEAASGARVLGGLLSKALLPTALLL 57
DQ 272 LLRPPHVALFTFDGLDELHDLSDLSRVPDSSC--PWEPAHPLVLLANLLSGKLLKASKL 329
QY 58 VITRAAAPGRLOGLRCSPOCAEVRGSDKDKKFKYKFFRDERRAERAYRVKVENETLPA 117
DQ 330 LTART---GIEVPRQFLRKVKLVLRGSPSHLRAYARMFPERALQDRLLSLEAFNLCS 386
QY 118 LCFVPFVCMIVCTV---LRQLELGRDLRSRTSTTTSVLLFTTSV-----L 161
DQ 387 LCSVPLFCWIIIFRCQHFRAAFEGSPQLPDCPTMTLTDVFL--VTEVHLNMQPSSLVQRN 445
QY 162 SSAPVADGPELQGLRLNCLAREGLVGRRAQPAEKELECLE-RGSKVOTLELSKELP- 220
DQ 446 TRSPVETLHAGRODTLCSLGGVAHGRGKSLFVFTQEVQASGLQERDMQJGFL--RALPE 503
QY 221 -GVLETEVTYQFIDQSFQEFIAALSYLEL-----GGVPRTAA----- 257
DQ 504 LGPGGDCQSYEFHLLTQAFFTAFFLVDDRVTGTOELLRRFFQEWMPAGAAATTCYPPFL 563
QY 258 -----GGVGTLLRGDAQPHSHLVLTTRFLFGLLSAERMRDIERHFGCMYSERVKQEARLW 312
DQ 564 PFQCLQSGGPAREDLFKNDHFQFTNLFLCGLLSKAKQK-LLRHLVPAALRRKRAL-W 621
QY 313 -----VOGQGGQCPGVAPEVTEGAKGLEDTBEPSEEEGEPNYPJELLYCLYETQED- 365
DQ 622 AHLFSSLRGVLKSLPRVQVESFNQVQAM-----PTF-IWMLRCIYETQSQK 666
QY 366 --AFVRAQALCRFPFELALQVRFCRMDVAVLSYVCRCCPAGQALRLISCRILVAQEKKKX 423
DQ 667 VGQLAARGICA-NYLKLTTCNACSADCSALSFVLHHPKRLALCLDNNL----- 715
QY 424 LGKELQASLGGSSQGTQKLPASLLHPLFQAMTDPDLCHLSLTLSSHCKLPDAVCLDSE 483
DQ 716 -----NDYGVR-----LQPCFSRLT-----VLRSLVNQITDGGVKVLSE 750
QY 484 ALRAAPALTELGLLHNRLEAGRLMSEGLAWPQCRVQTVRVCLPD---PQRCLOVY---- 537
DQ 751 ELTKYKIVTVLGLNQNQITDVGARYVTKIL--DECKGLT-HLKLGNKNTSEGKYLALA 807
QY 538 -----VGM-----LRQSPALTTLDLSCGLPAPMVTVLCAVLQHQ 572
DQ 808 VKNSKTSSEVGMGNQVGDGAKAFEAALRNHPSLTTLASNGISTEGGKS-LARALQ-Q 866
QY 573 GCGLOTLASVELSE---OSLOELQAVKR 599
DQ 867 NTSLEILWLTQNELNDEVAESLAEMLKVNQ 896
```

RESULT 5

US-09-340-620A-8

; Sequence 8, Application US/09340620A

; Patent No. 6482933

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED

; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF

; FILE REFERENCE: 07334-124001

; CURRENT APPLICATION NUMBER: US/09/340,620A

; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: US 09/245,281

; PRIOR FILING DATE: 1999-02-05

; PRIOR APPLICATION NUMBER: US 09/207,359

; PRIOR FILING DATE: 1998-12-08

RESULT 7
 US-09-207-359B-43
 : Sequence 43, Application US/09207359B
 : Patent No. 6469140
 : GENERAL INFORMATION:
 : APPLICANT: Bertin, John
 : TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
 : FILE REFERENCE: 07334-112001
 : CURRENT APPLICATION NUMBER: US/09/207,359B
 : CURRENT FILING DATE: 1998-12-08
 : PRIOR APPLICATION NUMBER: US 09/099,041
 : PRIOR FILING DATE: 1998-06-17
 : PRIOR APPLICATION NUMBER: US 09/019,942
 : PRIOR FILING DATE: 1998-02-06
 : NUMBER OF SEQ ID NOS: 47
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 43
 : LENGTH: 953
 : TYPE: PRT
 : ORGANISM: Mus musculus
 US-09-207-359B-43

Query Match 7.8%; Score 251.5; DB 4; Length 953;
 Best Local Similarity 22.2%; Pred. No. 8.8e-18;
 Matches 164; Conservative 87; Mismatches 298; Indels 191; Gaps 27;

QY	1	MLAQPORLLFIDGAEDELPAALGGPEAAP--CTDPFAASGARVLGGLSKALPTALLLV	58
DB	272	LLRFPHTALFTFDGLDELHSDFDLSRVPDSCC-PWEPAPHLVLANLLSGRLKAGKLL	330
QY	59	TTTAA--APGRLOGLRCLSPQCAEVRGFSKDKKKYFYKFFRDERRAERAYRVKNETLF	116
DB	331	TATGVEVPQLLRKK-----VLLRGFSPSHLRAYARMPEPTAQEHLLOQLDANKPLC	365
QY	117	ALCFVPFVCMIV--CTVLRQQLBELG--RDLRSRTSKTTTSVLLFTITSVLSSAPVADGPR	172
DB	386	SLCGVPFCWIIIFRCFQHFQTVFEGSSQLPDCAVLTIDVFLV-VTEVHLNRP-QFSSLV	443
QY	173	QGLDRLNCLRLAREG-----VLGRRAQ-----FAEKELEOLELRGSKVQTLFLSKKE	218
DB	444	QRNTRSPAETLRAGWRTLHALGEVAHRGTDKSLFVFGQEVQASKLQEGDLOLGFU--RA	501
QY	219	LPGV-LETEVYQFIDQSFQEFLLAALSYLLEDGGV-----PRTAA-----	257
DB	502	PDVGFEGQSGSYEFFHLTLQAFTAP-FLVADKVKSTRELLRFFREWTSPGATSSSCHS	560
QY	258	-----GGVGTLLRGDAQPHSHLVLTTRFLGLLSAERMRIERHFGCMVSERVQEA	309
DB	561	SFFSFCGLGGRSLRGLPDFFRNKDHQFTNLFVCGLLAKARQKLRLQVPAKILRR-KRKA	619
QY	310	LRVVOGQGCPCGVAPVETEGAKGLEDETEPEEBEGBEENPYLELLYCLYETQEDAFVR	369
DB	620	L-WAH-----LFASLSRYLSLPRVOSGGFNQVHAMPTF-LMWLRCIYETQSKVGR	669
QY	370	QALCRFPPELALQVRFCRMVDAVLSYCVRCPCAGQALRLISCLRVAAQEKKKSLGRLQ	429
DB	670	LA-----ARGISADYKLAFNACSADCSALSFVL-----HHFH	703
QY	430	ASLGGSSQGTTKQLPASLLHLPFOAMTDPCLHLSLTLSHCKLPDAVCRDSEALRAAP	489
DB	704	RLALDLDNNLNDYGVQELQPCFSRLT-----VIRLSVNIQITDGVKVLCEELTKYK	756
QY	490	ALTEGLLHRLNLSAELRMUSLSEGLAMPQCKVQTVRVQLPDPQGLQYL-----	537
DB	757	IVTFGLYNNQITDIGARYVA-----QILDECRGLKHLKGLGNRITSEGG	801
QY	538	-----VGM-----LRSPALTTLDLSGGQLPAPVYTYLC	566
DB	802	KVALAVKNTSIVDVGMGNQITGDECAFAFALKDHPSLTTLTSLAFNGISPEGGKSLA	861
QY	567	AVLQHOCSGQTLISLASVELSESQ-----ELQAVKEA	600

QY 49C ALTELGLHNRLEAGRLMSEGLAMPQCRVQTVRVLQDPQOGLQYL----- 537
DB 757 IVTFGLYNNQITDVGARYVA-----CILDECRGJXHLKGXNRITSEGG 802
QY 538 -----VGM-----LROSPALITLDSGQCLPAPMVTYLC 566
DB 802 KVALAVKXSTSVIVGMXGNCIGDEGAKAFABALKDHPSLTSLAFNGISPEGCKSLA 861
QY 567 AVLQHQCGGLQTLASVLSBESQSLQ-----ELQAVKRA 600
DB 862 QALK-QNTLTVLWLTQNEJNDVGMNQVGEAGAKAFABALKDHPSLTSLAFNGISPEGCKSLA 920
QY 601 KPDJ-VTHRALDGHPPQPE 620
DB 921 QXNTA-TE-CLNGNJKPEE 940

RESULT 9
US-09-207-359B-47
; Sequence 47, Application US/09207359B
; Patent No. 6469140
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-112001
; CURRENT APPLICATION NUMBER: US/09/207,359B
; PRIOR FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 966
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: majority sequence
US-09-207-359B-47

Query Match 7.0%; Score 224.5; DB 4; Length 966;
Best Local Similarity 23.1%; Pred. No. 7.9e-15;
Matches 161; Conservative 84; Mismatches 296; Indels 155; Gaps 30;

QY 1 MLAGPQRLFLFDGADLPA---LGGPEAACTDPFEAASGARVLGGLSKALLPTALL 57
DB 272 LARFPHVALFTDGLDELHSDLSRVPDSSC--PWEPAHPLVLLANLLSGLLKAGKL 329
QY 58 VTTBAA--APGR-OGRLCSQCAEVGFSKDKKKYKFFRERRAERAYRVKNETJ 115
DB 330 LTARTGVEVPRQLRKK-----VLLRGFSSHRLAYARRFPERRAQQDLSCIDANPL 384
QY 116 FALCFVPFVCMIV--CTVLRLQLELG--RDLRSRTSKTTTSWLL-----FITSVLSAPVA 167
DB 385 CSJCGPVLFWIIFRCQHQFAAFEGSSSLQPDCAVTLTDVFLVTEVHLNRNCPSSLVQ 444
QY 168 DGRFLQGD-----LRNLCRIAREGVLRRAQFAKELEQLELGSKVQTLFLSKKEL 219
DB 445 RNTRSPAETLHAGRTLHAGVAHRGTDSKSLFVFGOEYQASGLQEGDLQLGFL--RAL 502
QY 220 P--GVLETEVYQIDOSFQEFLLAALSYLLED-----GGVPRTA--- 257
DB 503 PWDGPGGGGSGYEFHUTLQAF--AFFLVADCKVGTQELJRRFFQWMTSPGGAASSCHSS 562
QY 258 -----GGVGTLLRGDAQFHSVLVTLTFLGLLSAERMRDIERHFGCMVSRVKOEAL 310
DB 563 FLSFQCLGSGRAGEDJ-FKVKQHFQFTNLFVCGLL-AKAKQKLRQLVPAALRRKRAL 621
QY 311 RWVQ--QGQSGCGVAVEVTEGAKGLEDTPEEBEEGEEPNYPLELLYCLYETQEDAFV 368

DB 622 -WAHLFASLRYGLKSLPRVQVG--GFNQVQ-----AMPTF-LWMLRCIYE-TSQO--- 666
QY 369 ROALCRFPFELALQVRFCRMDVAVLSYVCRCCPAGQALRLISCRLEVAQEKKKSLGKRL 428
DB 667 -----KVGOLA---ARGISADYKLAFCNACSDACSLSFV-----LHFFHKQL 707
QY 429 QASLGGSSQGTTKQLPASLILHPJFOAMTDPFLCHLSLTLSSHCKLPDVCARDJSEALRAA 488
DB 708 ALDJ-----DNNLNDYGVQELQPCFSLT-----VRLSVNQITDGGVXVJSEELTKY 756
QY 489 PALTELGLHNRLEAGRLMSEGLAMPQCR-----VQTVRVQLPD-----PQGL 534
DB 757 KIVTFGLYNNQITDVGARYVAQIL--DECKGLHLSLYNNQITDVGAKLKXKITSEGG 814
QY 535 QYL-----VGM-----LROSPALITLDSGQCLPAPMVTYLC 566
DB 815 KYVALAVKXSTSVIVGMXGNCIGDEGAKAFABALKDHPSLTSLAFNGISPEGCKSLA 874
QY 567 AVLQHQCGGLQTLASVLSBESQSLQ-----ELQAVKRA 599
DB 875 QALK-QNTLTVLWLTQNEJNDVGMNQVGEAGAKAFABALKDHPSLTSLAFNGISPEGCKSLA 909

RESULT 10
US-08-910-731-2
; Sequence 2, Application US/08910731;
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: CS/08/910,731
; FILING DATE: (Herewith)
; PRIOR APPLICATION DATA:
; CLASSIFICATION: 435
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942,3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-731-2

Query Match 6.7%; Score 216; DB 2; Length 456;

Best Local Similarity 27.4%; Pred. No. 1.9e-14;
Matches 77; Conservative 45; Mismatches 113; Indels 46; Gaps 7;
QY 354 ELLYCLYETQ----EAFVRCALCRFPPELAQVRFCRMYAVLSYCVRCPCAG-----403
DB 17 ELLPLQQVEVRLDDCGLTEHCKDIGSAL-----RANPSLTELCLRTNELGDAGVHL 70
QY 404 --QALRLSCRIVAAQEKKKSLGRKLQASLGGSOGTTKQLPA-----SL 448
DB 71 VLQGSQPTCKI-----QKLSLQNCSTLTAAGCGVLPSTLRSLPTRELHLSNDPLGDAG 124
QY 449 LHPFLQAMTDPCLHSLSLTLSHCKLPDAVCRDLSEALRAAPALTELGLLHNLSEAGLRM 508
DB 125 LRLCEGLDPPCHLEKLEQLEVCRLTAASCEPLASVLRATRAKELTVSNNDIGEAGARV 184
QY 509 LSEGAWPQCRVQTVRVQLPDPQRGL-----QYLVMRLRQSPALTTLDLSCQPLPAPMWT 563
DB 185 LQGLGADSAQOLETIRLE-----NCGLTTPANCKDLGCIIVASQASRLRELDLGSNGLGDAGIA 240
QY 564 YCAVLQHCGCGLOTLTLASVLSLSEQSLQELQAVKRAKPD 604
DB 241 ELCPLGLSPASRLKTLWLWECDITASGCRDLRCRVLOAKETL 281

RESULT 11
US-08-795-395-2
; Sequence 2, Application US/08795335
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-795-395-2

Query Match 6.7%; Score 216; DB 2; Length 456;
Best Local Similarity 27.4%; Pred. No. 1.9e-14;
Matches 77; Conservative 45; Mismatches 113; Indels 46; Gaps 7;
QY 354 ELLYCLYETQ----EAFVRCALCRFPPELAQVRFCRMYAVLSYCVRCPCAG-----403

DB 17 ELLPLQQVEVRLDDCGLTEHCKDIGSAL-----RANPSLTELCLRTNELGDAGVHL 70
QY 404 --QALRLSCRIVAAQEKKKSLGRKLQASLGGSOGTTKQLPA-----SL 448
DB 71 VLQGSQPTCKI-----QKLSLQNCSTLTAAGCGVLPSTLRSLPTRELHLSNDPLGDAG 124
QY 449 LHPFLQAMTDPCLHSLSLTLSHCKLPDAVCRDLSEALRAAPALTELGLLHNLSEAGLRM 508
DB 125 LRLCEGLDPPCHLEKLEQLEVCRLTAASCEPLASVLRATRAKELTVSNNDIGEAGARV 184
QY 509 LSEGAWPQCRVQTVRVQLPDPQRGL-----QYLVMRLRQSPALTTLDLSCQPLPAPMWT 563
DB 185 LQGLGADSAQOLETIRLE-----NCGLTTPANCKDLGCIIVASQASRLRELDLGSNGLGDAGIA 240
QY 564 YCAVLQHCGCGLOTLTLASVLSLSEQSLQELQAVKRAKPD 604
DB 241 ELCPLGLSPASRLKTLWLWECDITASGCRDLRCRVLOAKETL 281

RESULT 12
US-08-910-731-8
; Sequence 8, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith);
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 38/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-910-731-8

Query Match 6.7%; Score 214; DB 2; Length 456;
Best Local Similarity 27.4%; Pred. No. 3.1e-14;
Matches 77; Conservative 45; Mismatches 113; Indels 46; Gaps 7;

QY 354 ELIYCIYETQ---EDAFYRCALCRPELALQVRFCRMJVA'LSYCVCECCPAG-----4C3
DB 17 ELIPLLIQVEVRLDDCGTEHCKDIGSAL-----RANPSITEJCLRTNIGLGDAGVEL 7C
QY 4C4 --QALRLISCRVAAGEKKKISLKGRLQASLGGSSQGTTKQLPA-----SL 443
DB 71 VLOGLOSPTCKI-----QKLSLONCSLITAGCGVLPSTLRSIPTPELHLSNPLGDAG 124
QY 449 LHPFQAMTDPJCHLSJLTSJSHCKJGPDVAVCRDISEALRAAPALTEJGLHNLRLSAGLKM 5C8
DB 125 LRLLCGLLDPOCHLEKQJFYCRJTAASCEPLASVLRATRAKELTVSNNDIGEAGARV 184
QY 509 LSEGLAWPOCRVCTVQJPDPRGJ-----QVLVGLRQSPALTTLDLSGCOLPAPMTV 563
DB 185 LGGLGADSAQOLETLRZE-----NCGJTPANCKDLGIVASQASRLSRLALGSKLGDVGMA 24C
QY 564 YLCVAVLQHCQGLQTLASVLSVSEOSLQELQAVKSAKFDL 604
DB 241 ELCPGLLHPSRLRLTLWIWEGGITAKGCGDLORVLRKESL 281

RESULT 13

US-08-910-731-6
; Sequence 6, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.344C003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide

US-08-910-731-6

Query Match: 6.6%; Score 212; DB 2; Length 461;

Best Local Similarity 29.2%; Pred. No. 5.2e-14;
Matches 59; Conservative 35; Mismatches 104; Indels 4; Gaps 2;
QY 404 QALRLISCRVAAGEKKKISLKGRLQASLGGSSQGTTKQLPASLHPLFQAMTDPJCHL 463
DB 31 QVRLDDCGLT---EARCKDISSALRVNPAALAEINLRSNELGDVGHCVLQGLQTPSCKI 87
QY 464 SSLTSHCKLPDANCRDISEALRAAPALTEJGLHNLRLSAGLKMSEGLAWPOCRVQTV 523
DB 88 QKLSZONCCLTGAGCGVLSSTLRTPLTQELHLSNLLGDAGLQGLLDPQCRLEKL 147
QY 524 RVGJPD-PQRGLQVLMRLSPALTTLDLSGCOLPAPMTVYLCVAVLQHCQGLQTL-SLA 582
DB 148 QLEVCSLSAASCEPLASVLRAPKDFKELTVSNNDINEAGVRVLCQGLXDSFCQLEAKLE 2C7
QY 583 SVLSVSEOSLQELQAVKRAKPDJ 604
DB 228 SCGVTSDNCRDLGIVASKASL 229

RESULT 14

US-08-910-731-4
; Sequence 4, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:

APPLICANT: CHATTERJEE, DEB K.
APPLICANT: SHANDILYA, HARINI
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.344C003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-910-731-4

Query Match: 6.5%; Score 209.5; DB 2; Length 456;

Best Local Similarity 24.3%; Pred. No. 9.6e-14;
Matches 93; Conservative 45; Mismatches 129; Indels 115; Gaps 15;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: October 2, 2003, 17:35:58 ; Search time 24 Seconds
(without alignments)
1224.654 Million cell updates/sec

Title: US-10-029-347-2

Perfect score: 3218
Sequence: 1 MLAPQORLLPILGADLPA.....ITHPALDGHPPKELISTF 625

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	IC	Description
1	3207.5	99.7	892	1	PYAS_HUMAN		P59544 homo sapien
2	2183	67.8	854	1	PYAS_RAT		G63315 rattus norv
3	227.5	66.1	843	1	PYAS_MOUSE		G91452 mus musculus
4	835	25.9	1033	1	CISI_MOUSE		C8r4b8 mus musculus
5	825	25.6	1062	1	PVAF_HUMAN		P59046 homo sapien
6	800.5	24.9	1034	1	CISI_HUMAN		G96D20 homo sapien
7	762	23.7	1473	1	NALI_HUMAN		G3C000 homo sapien
8	650	20.2	980	1	PVAF_HUMAN		Q8wx94 homo sapien
9	616.5	19.2	1062	1	NALI2_HUMAN		G9nxd2 homo sapien
10	598.5	18.6	994	1	NALI4_HUMAN		G96m22 homo sapien
11	507.5	15.8	1200	1	MATE_HUMAN		P59047 homo sapien
12	468.5	14.6	1111	1	MATE_MOUSE		G9r1m5 mus musculus
13	461	14.3	1033	1	PVAF_HUMAN		P59045 homo sapien
14	305	9.5	1020	1	CARF_MOUSE		G8k320 mus musculus
15	304.5	9.5	1130	1	C2TA_HUMAN		P33076 homo sapien
16	300.5	9.3	1040	1	C2TA_HUMAN		G9hc29 homo sapien
17	289	9.0	953	1	CAR4_HUMAN		G9y239 homo sapien
18	267.5	8.3	1155	1	C2TA_MOUSE		P79621 mus musculus
19	254.5	7.9	953	1	CAR4_MOUSE		G8bhb0 mus musculus
20	216	6.7	456	1	RINI_PIG		P10775 sus scrofa
21	212	6.6	460	1	RINI_HUMAN		P13489 homo sapien
22	209.5	6.5	456	1	RINI_RAT		P29315 rattus norv
23	123	3.8	1024	1	CARC_HUMAN		G9npp4 homo sapien
24	116	3.6	1403	1	BIRE_MOUSE		G9r016 mus musculus
25	113.5	3.5	681	1	SSAV_SALTY		P74856 salmonella
26	112.5	3.5	1238	1	SRCC_RHOCA		G68032 rhodobacter
27	108.5	3.4	406	1	HUTI_VIBRIO		Q8da18 vibrio vuln
28	107.5	3.3	780	1	PRTP_HSVIF		P06490 herpes simp
29	107.5	3.3	3321	1	PCN2_HUMAN		G95613 homo sapien
30	106	3.3	1173	1	UB4B_MOUSE		G9es00 mus musculus
31	104.5	3.2	1795	1	ES21_HUMAN		Q14674 homo sapien
32	104	3.2	857	1	CA2P_THES7		P51060 thermus sp.
33	104	3.2	1403	1	BIRA_MOUSE		G9qwk5 mus musculus

RESULT 1

ID	PYAS_HUMAN	STANDARD;	PRT;	892 AA.
AC	P59044;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	PYRIN-containing APAF1-like protein 5.			
GN	PYPAF5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
CX	MCS1_TaxID=9606;			
RY	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22162427; PubMed=12019269;			
RA	Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,			
RA	Lora J.M., Geddes B.J., Briskin M., Disteferano P.S., Bertin J.			
RT	"PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates			
RT	activation of NF-kappa B and caspase-1-dependent cytokine			
RT	processing."			
RL	J. Biol. Chem. 277:29874-29880(2002).			
RY	[2]			
RP	FUNCTION:			
RX	MEDLINE=22275822; PubMed=12387869;			
RA	Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,			
RA	Kelly R., Carlson A., Merriam S., Lora J.M., Briskin M.,			
RA	DiStefano P.S., Bertin J.			
RT	"Functional screening of five PYPAF family members identifies PYPAF5			
RT	as a novel regulator of NF-kappa B and caspase-1."			
RL	FEB Lett. 530:73-78(2002).			
CC	-!- FUNCTION: May mediate activation of CASP1 via ASC and promote			
CC	activation of NF-kappa-B.			
CC	-!- SUBUNIT: Binds to ASC with its DAPIN domain.			
CC	-!- SIMILARITY: Contains 1 DAPIN domain.			
CC	-!- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at			
CC	much lower levels in T-cells.			
CC	-!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.			
CC	-!- SIMILARITY: Contains 1 NACHT domain.			
CC	-!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AF479748; AA487105.1; -			
DR	InterPro; IPR007091; LRR_RNinh.			
DR	InterPro; IPR003590; LRR_RNinh.sub.			
DR	InterPro; IPR007111; NACHT_NTPase.			
DR	InterPro; IPR004020; PAAD_DAPIN_dom.			
DR	Pfam; PF02758; PAAD_DAPIN; 1.			
DR	SMART; SM00388; LRR_R1; 3.			

ALIGNMENTS

```

DR PROSITE; PS50824; DAPIN; 1.
KW PROSITE; PS50837; NACHT; 1.
FT DOMAIN; 1 103 DAPIN.
FT DOMAIN; 156 513 NACHT.
FT REPEAT; 604 614 POLY-GLU.
FT REPEAT; 462 487 LRR 1.
FT REPEAT; 727 747 LRR 2.
FT REPEAT; 755 778 LRR 3.
FT REPEAT; 811 834 LRR 4.
FT REPEAT; 945 868 LRR 5.
FT NP_BIND; 202 209 ATP (POTENTIAL).
SQ SEQUENCE 892 AA; 95733 MW; 4AABDFC766DDE9D CRC64;

Query Match 95.7%; Score 3207.5; DB 1; Length 892;
Best Local Similarity 95.8%; Pred. No. 1.3e-226;
Matches 625; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Cy 1 MLAGPQRLFLDGADELPAALGPEAAFCDDPEAASGARVLGGLSKALPTALLVIT 60
D5 267 MLAGPQRLFLDGADELPAALGPEAAFCDDPEAASGARVLGGLSKALPTALLVIT 326
Cy 61 RAAPAGRLQRLCSQCAEVGSPDKKKYFYFFDERRAERAYRFVKNETIFALCF 120
D5 327 RAAPAGRLQRLCSQCAEVGSPDKKKYFYFFDERRAERAYRFVKNETIFALCF 386
Cy 121 VPFVCMIVCTVLRQGLRGDLSRTSKTTSVLLFITSVLSAPVADGRLQDLNLC 180
D5 367 VPFVCMIVCTVLRQGLRGDLSRTSKTTSVLLFITSVLSAPVADGRLQDLNLC 446
Cy 181 RLAREGVLRRAQPAEKELELRGSKVQTLFLSKKELFGVLETEVTFIQDSQOEFL 240
D5 447 RLAREGVLRRAQPAEKELELRGSKVQTLFLSKKELFGVLETEVTFIQDSQOEFL 506
Cy 241 AALSYLEDDGVPRTAGGVJTLRGDAQPHSHLVLTTRFLGILLSAERNRDIERHFGCM 300
D5 507 AALSYLEDDGVPRTAGGVJTLRGDAQPHSHLVLTTRFLGILLSAERNRDIERHFGCM 566
Cy 301 VSERVKQEARLWVGOGGCGVAPETEGAKGLDTEEPREEEGEPNYPLELLYCLY 360
D5 567 VSERVKQEARLWVGOGGCGVAPETEGAKGLDTEEPREEEGEPNYPLELLYCLY 626
Cy 361 ETQDAFVRAQALCFPELALQVRFCRMDVAVLSYCVRCPPAGQALRLISCLVAQAEKK 420
D5 627 ETQDAFVRAQALCFPELALQVRFCRMDVAVLSYCVRCPPAGQALRLISCLVAQAEKK 686
Cy 421 KKSIGKRLQASLGGG-SQGGTTKQLPASLLHPLFQAMTDPLCHLSLTLGHCKLPDAVCR 479
D5 687 KKSIGKRLQASLGGGSSQGGTTKQLPASLLHPLFQAMTDPLCHLSLTLGHCKLPDAVCR 746
Cy 480 DLSBALRAAPALTELGLLHNLRLSAGLRLMSEGLAWPQCRVQTVRVOLPDQSGLOYLVG 539
D5 747 DLSBALRAAPALTELGLLHNLRLSAGLRLMSEGLAWPQCRVQTVRVOLPDQSGLOYLVG 806
Cy 540 MLRQSPALTTLDLSGCOLPAPMTVYLCVILQHQSCGLQT-SLASVELSEQSLQELQAVKR 599
D5 807 MLRQSPALTTLDLSGCOLPAPMTVYLCVILQHQSCGLQT-SLASVELSEQSLQELQAVKR 866
Cy 600 AKPLVITTHALDGHGPPKELISTF 625
D5 867 AKPLVITTHALDGHGPPKELISTF 892

RESULT 2
PYAS-RAT
ID PYAS5 RAT STANDARD; PR7; 854 AA.
AC Q63C35;
DC 28-FEB-2003 (Rel. 41; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DE 15-SEP-2003 (Rel. 42; Last annotation update)
DE PYRIN-containing APAF1-like protein 5-like (Angiotensin II/vasopressin receptor).
GN PYPAF5 OR AVR.

```

```

CS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP CONCEPTUAL TRANSLATION OF 1-343.
RA Hinz U.;
RN Unpublished observations (FEB-2003).
RN [2]
SEQUENCE OF 344-854 FROM N.A., MUTAGENESIS OF GLU-393, AND FUNCTION.
RC TISSUE=Kidney;
RX MEDLINE=21980185; PubMed=7489366;
RA Ruiz-Opazo N., Lopez L.V., Herrera V.L.M.;
RT "Identification of a novel dual angiotensin II/vasopressin receptor on
RT the basis of molecular recognition theory.";
RG Nat. Med. 1:1074-1081(1995).
RN [3]
RP VARIANTS SALT-SENSITIVE HYPERTENSION SER-492 AND ARG-536.
RX MEDLINE=21980185; PubMed=11984003;
RA Ruiz-Opazo N., Lopez L.V., Herrera V.L.M.;
RT "The dual AngII/AVP receptor gene N119S/C163R variant exhibits
RT sodium-induced dysfunction and cosegregates with salt-sensitive
RT hypertension in the Dahl salt-sensitive hypertensive rat model.";
RG Mol. Med. 8:24-32(2002).
RN [4]
RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
RT "Identification of mammalian orthologs associated PYPAF5 with distinct
RT functional roles.";
RG FEBS Lett. 538:173-177(2003).
RN [5]
CC -!- FUNCTION: May mediate activation of CASP: via ASC and promote
CC activation of NF-kappa-B (By similarity). Angiotensin II and
CC vasopressin binding protein. May stimulate cAMP accumulation.
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney. Detected at low
CC levels in all tissues tested.
CC -!- DISEASE: Defects in PYPAF5 may be a cause of salt-sensitive
CC hypertension.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
CC -!- CAUTION: The N-terminus was extended using ESTs and genomic
CC sequences, in analogy to ortholog sequences.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M85183; AAA03623.1; ALT_INIT.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR Pfam; PF00560; LRR; 2.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 103 DAPIN.
FT DOMAIN 168 484 NACHT.
FT DOMAIN 564 604 ASP/GLU-RICH.
FT DOMAIN 655 662 POLY-LYS.
FT REPEAT 433 458 LRR 1.
FT REPEAT 610 633 LRR 2.
FT REPEAT 722 745 LRR 3.
FT REPEAT 812 836 LRR 4.
FT NP_BIND 174 181 ATP (POTENTIAL).
FT VARIANT 492 492 N -> S (IN SALT-SENSITIVE HYPERTENSION).
FT VARIANT 536 536 C -> R (IN SALT-SENSITIVE HYPERTENSION).
FT MUTAGEN 766 766 E->K; ABOLISHES ANGIOTENSIN II BINDING.

```



```

SQ SEQUENCE 854 AA: 95292 MW: D7BBD922D2D77B734 CRC64;
Query Match 67.8%; Score 2183; DB 1; Length 854;
Best Local Similarity 71.0%; Pred. No. 8,6e-152;
Matches 443; Conservative 55; Mismatches 112; Indels 14; Gaps 5;
QY 1 MIAQPORLLFILDGADLPALGPEAAPCTDPEAASGARVLGGLLSKALJPTALLVTT 60
DB 239 ILAQPHRLFLDGADELPTLAAPEATPCRDPEATSGURVLSGLLSQELJFSARELVTS 298
QY 61 RAAAPGRLGRLCSPCAEVEGFSKDKKKYKFFRDERRAERAYRFVKEVETLFCALCF 120
DB 299 RNATLGRQLGRLCSPCAEVEGFSKDKKKYKFFRDERKAERAYRFVKEVETLFCALCF 358
QY 121 VPFCVTCVTLVQOELGRLGRLSRTSKTTTSVYLLFITSVLSAPVADPRLGRLNLC 180
DB 359 VPFCVTCVTLVQOELGRLGRLSRTSKTTTSVYLLFITSMLKSAG-TNPRVQSGRLNLC 417
QY 181 RLAREGVLRRAQFAEKELEOLELRGSKVQTLFLSKKELPGVLETVTYQFIDQSFQBL 240
DB 418 RLAREGILKHAQFSEKLEFLKLGQSQVQVTFLSKKELPGVLETVTYQFIDQSFQBL 477
QY 241 AALSYLLEDCGVPRTAGGVTLLRDAQPHSHVLTTRFLGLLSAERMRIERHFGCM 300
DB 478 AALSYLLEDCGVPRTAGGVTLLRDAQPHSHVLTTRFLGLLSAERMRIERHFGCM 537
QY 301 VSERVKEALRWVGGGQCGVA---PEVTGAKGLEDTESPEREEGEERPNYLELLY 357
DB 538 VGRVVKQDTLRVVGQSGQ--PKVATVGAKKDELKDEAESEEEEEENFGLELLY 595
QY 358 CLVETQEDAFVQALCRFPALQVRFCRMVAVSYCVRCPCAGQALRLSCRLVAAQ 417
DB 596 CLVETQEDAFVQALCRFPALQVRFCRMVAVSYCVRCPCAGQALRLSCRLVAAQ 655
QY 418 E--KKKSLGKRLQASLGGSSGQTKQLPASLHPFOAMTDPLCHSLSLTSHCKLPD 475
DB 656 EKKKKKKFMNRJL-----GSOSTGQPPASLRPLCEATQCGSLTSLHCKLPD 709
QY 476 AVCRLDLSEALRAAPALTEJGLHNRJSEAGRLMSEGLAWPQCRVQTVRVQLPDPQRGQ 535
DB 710 AVCRLDLSEALRAAPALTEJGLHNRJSEAGRLMSEGLAWPQCRVQTVRVQLPDPQRGQ 595
QY 536 YLVGMRLQSPALTTDLSSGQCPAPVWTVYLCVQLHQGGLQTLASVSESEQSLQELQ 595
DB 770 YLVGMRLQSPALTTDLSSGQCPAPVWTVYLCVQLHQGGLQTLASVSESEQSLQELQ 829
QY 596 AVKRAKPDVITHPALDGHGCPBPX 619
DB 830 AVKRAKPDVITHPALDGHGCPBPX 853
RESULT 3
PYA5 MOUSE STANDARD; PRT; 843 AA.
AC G91ME2; O8KO24;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PYRN-containing APAFI-like protein 5-like.
GN PYPAF5.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP CONCEPTUAL TRANSLATION OF 1-287.
RA Hinz U.;
RL Unpublished observations (FEB-2003).
RN (2)
RP SEQUENCE OF 288-843 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,

```

```

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschui S.F., Zeeberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Xie S.L., Wang C., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalun D.X., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield J.S.N., Krzywinski M.I., Skalek U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[3]
RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
RA "Identification of mammalian orthologs associates PYPAF5 with distinct
RT functional roles."
RL FEBS Lett. 538:173-177(2003).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC activation of NF-kappa-B (By similarity).
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.
CC -!- CAUTION: The N-terminus was extended using ESTs and genomic
CC sequences, in analogy to ortholog sequences.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EXBL: BC013519; AAH13519.1;
CC EMBL: BC031139; AAH31139.1; ALT_INIT.
CC MGD: MGI:2141990; Pypaf5.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR007091; LRR_RNinh.
CC Pfam: PF00560; LRR; 1.
CC PROSITE: PS00824; DAPIN; 1.
CC PROSITE: PS00837; NACHT; 1.
CC KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 11 102 DAPIN.
FT DOMAIN 168 485 NACHT.
FT DOMAIN 569 585 ASP/GJU-RICH.
FT DOMAIN 654 661 POLY-LYS.
FT REPEAT 434 459 LRR 1.
FT REPEAT 609 632 LRR 2.
FT REPEAT 811 834 LRR 3.
FT NP_BIND 174 181 ATP (POTENTIAL).
SQ SEQUENCE 843 AA; 94592 MW; 35FB7A766A47DB51 CRC64;
Query Match 66.1%; Score 2127.5; DB 1; Length 843;
Best Local Similarity 70.2%; Pred. No. 9,6e-148;
Matches 431; Conservative 63; Mismatches 103; Indels 17; Gaps 7;
QY 1 MIAQPORLLFILDGADLPALGPEAAPCTDPEAASGARVLGGLLSKALLPTALLVTT 60
DB 239 ILAQPHRLFLDGADELPTLPSSEATPCPKDPLEATSGRLVLSGLLSQELPCGRLVTT 298
QY 61 RAAAPGRLGRLCSPCAEVEGFSKDKKKYKFFRDERRAERAYRFVKEVETLFCALCF 120
DB 299 RHAATGRLGRLCSPCAEVEGFSKDKKKYKFFRDERKAERAYRFVKEVETLFCALCF 358

```

```
QY 121 VPVCMVCTVLPQOELGHD:SRSTKTTTSVLLFITSVLSSAPVADGPRQGLDNLIC 180
DB 359 VPVCMVCTVLPQOELGHD:SRSTKTTTSVLLFITSVLSSAPVADGPRQGLDNLIC 417
QY 181 RLAREGVJG-RAAQAFAEKELEQLELRGSKVQTLFLSKKEIPGVJETEVITYCFIDQSFQEF 239
DB 418 RLAREGILDHKKAFQSEEDLEKLRGSOVQTIPLNKEIPGVLTKEVITYCFIDQSFQEF 477
QY 240 LAALSYLEDDGVPRTAAGVGTLLRGDAQPHSHLVLTTRFLRGLLSAERMRDIERHFQC 239
DB 478 LAALSYLEAERTPGTPAGGVQKLLNSDAELRGLAATTRFLRGLLNTGCRDIGNHFQC 537
QY 303 VVSERVKQEARLVVQGG--GGCGVAPVETGAKG:EDTEEPDEEEEGEP-NYPLELL 356
DB 538 VPDHVKQOTIRVQGGSHPKGPPVGAKTAE-----LEDIEDAEEEEEEDUNFLELL 593
QY 357 YCYETQEDAFVRQALCRFPPELALQVRFCRDVAVLSYCVRCPCAGQALRL:SCRLVAA 416
DB 594 YCYETQEDAFVRQALSSLEIPLVRLTRMDLVNLVYCVQCPDQALRLVSCG:VAA 653
QY 417 QE---KKKSLGKRLQASLGSSGGTTKQLPASLHPPLFOAMTDPLCHLSLTLSHCKLP 474
DB 654 KKKKKKKSLVXGLK-----GQSTKKQPPVSLRLPLCTMTIPKCHLSVLL:LSHCRLP 707
QY 475 DAVCRDLSALRAAPALTEGLL:HNRLSEAGRLVSLSEGLAWPCQRYCTVRYQLPDPQRL 534
DB 708 DAVCRDLSALRAAPALTEGLL:HNRLSEAGRLVSLSEGLAWPCQRYCTVRYQLPDPQRL 534
QY 535 QYLVGMRLQSPALTTLDLSCQLPAPMYTLCVAVLQHQGGGLQTLSLASVELSEQL 594
DB 768 NYLVIVVQSPVLTTLTDLSCQLPAPMYTLCVAVLQHQGGGLQTLSLASVELSEQL 827
QY 595 OAVKRAKPLVITH 608
DB 828 QAVKTSKPLSLIY 841

RESULT 4
CIS_MOUSE STANDARD; PRT; 1033 AA.
AC QBR4B5;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cold autoinflammatory syndrome 1 protein homolog (PYRIN-containing
GN APAF1-like protein 1) (Mast cell maturation inducible protein 1).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN NCBI_TaxID=10090;
RP [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/cJ;
RA Kikuchi-Yanoshiba R., Koga K., Takeuchi Y., Sugiki T., Saito T.,
RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;
RT "Identification of inducible genes during in vitro maturation of mouse
RL bone marrow-derived mast cells to connective tissue-type mast cells."
CC Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May function as a potential inducer of apoptosis.
CC Interacts selectively with apoptosis-associated specklike protein
CC containing a CARD domain (ASC). This complex may function as an
CC upstream activator of NF-kappaB signaling (By similarity).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL cutation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ies-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB SMDL; AF486632; AAL90874.1; -
DR MGD; MGI:2653833; Cias1.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR007111; NACHT_NTFase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SMO0368; LRR_R1; 1.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00837; NACHT; 1.
KW Apoptosis; Repeat; Leucine-rich repeat.
FT DOMAIN 1 91
FT 216 532 NACHT.
FT REPEAT 737 760 LRR 1.
FT REPEAT 794 817 LRR 2.
FT REPEAT 851 874 LRR 3.
FT REPEAT 880 903 LRR 4.
FT REPEAT 908 931 LRR 5.
FT REPEAT 937 964 LRR 6.
FT REPEAT 965 988 LRR 7.
SQ SEQUENCE 1033 AA; 118274 MW; 5924690966812117 CRC64;

Query Match 25.9%; Score 835; DB 1; Length 1033;
Best Local Similarity 29.13%; Pred.No.3.1e-53;
Matches 226; Conservative 112; Mismatches 239; Indels 134; Gaps 15;

QY 1 MAAQORQLFLIDGADQLP-ALGGPEAAAPTCTDPPEAASGARV:GGLLSKALLPTALLVLT 59
DB 286 ILRKESRILFLMDGFDELQAFDEHIGEVCTDWKAVRGDILLSSLRKLLPKASLIIT 345
QY 60 TAAAPGELQORLCSQCAEVGRFSKDKKKYFYKFFDRERRAERAYFVENETLFLALC 119
DB 346 TRPVALEKLQHLDPHVEILGFSEAKRKEYFFKYFSNLQAREAPRLIQENEVLFTMC 405
QY 120 FVPFVCMVCTVLPQOELGHD:SRSTKTTTSVLLFITSVLSSAPVADGPRQGLDNLIC 179
DB 406 FIPVCMVCTVLPQOELGHD:SRSTKTTTSVLLFITSVLSSAPVADGPRQGLDNLIC 465
QY 180 RLAREGVJG-RAAQAFAEKELEQLELRGSKVQTLFLSKKEIPGVJETEVITYCFIDQSFQEF 239
DB 466 CSLAAGDIGNWQKILFEEDLRLGHGLOKTDV-SAF-RMNVFQKEVDCERFYFSHMTFQEF 524
QY 240 LAALSYLEDDGVPRTAAGVGTLLRGDAQ-----PHSHLVLTTRFLRGLLSA 287
DB 525 FAAMYLLLEAEAEGETVRKPGGCSLLNRDVKVLLNENYKGFEGYLFVVRFLFGVYVQ 584
QY 288 ERMRDIERHFQCMVSRVQKQALRWVQGGCGPGVAPETGAKGLEDTEPEEEEGE 347
DB 585 ERTSYLEKLSCKISQVRLELLKWI-----EVKAKAKLQ--WQPSO----- 625
QY 348 EPNYPLELLYCLYETQEDAFVRQALCRFPPELALQVRFCRDVAVLSYCVRCPCAGQALRL 407
DB 626 -----LELFYCY:YENQEEEDFVQSAMDFHFKIEIN--LSTRMDHVVSFSCIKNCHRVKTL 678
QY 408 LLSCLRVAAQKKKSLQKRLQ-----ASLGGG--SQQGTTK 442
DB 679 LGFFHNSPKKEEERGRGRPLDQVOCVFPDTHVACSSRLVNCCLTSSFCRGLFSSLSNLR 738
QY 443 QJ-----PASLLHPLFOAMTDPLCHLSLTLSHCKLPDAVCRDLSALRAAPA 490
DB 739 SUTELDLSNTLGDGFG--NRVLCALQHPGNCIQRLWLRGCGLSHQCCFDISSVLSQQK 796
QY 491 LTELGLLHNRLESEAGRLMSLEGLAWPCQRYCTVRYQLPDPQRL 519
DB 797 LVLELDLSNALGDGFGRLCLCVGLHLLCNLQKLVLSCLTSACCQDLALV:SSNHLSTR 856
QY 520 -----VQTVRVLDPDQGLQYL----- 537
```

```

Db 857 LYIGENALGDSGVQLCEKXKPCQCN:QKLGYNLSGLTSCCSALTSVLKTNQNFTHLYL 916
QY 538 -----VGMRLQSPALTTLDLGGCQLPAP-----560
Db 917 RSNALGDTGLRLCEGLHPDCKLQMLELONCSLTSRSCWN:STLIITHNSLRKMLGN 976
QY 561 ----WVLYCAVLQHQGCG:CTLASLASVLSLQSLQELQAVKRAKPDVLI 606
Db 977 DLGLDLCVTLCEVLKQKQGLQSLQELQELGEMYNLREITRALEALQERPELT 1027

RESULT 5
PYA7 HUMAN
ID PYA7 HUMAN STANDARD; PRT; 1062 AA.
AC PS9046;
DT 28-FEB-2003 (Rel. 41, Created;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PYRIN-containing APAF1-like protein 7 (March-1).
GN PYPAF7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Xamalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1].
RP MEDLINE=22162427; PubMed=12019269;
RA Wang J., Yanji G.A., Grenier J.M., Al-Carawi A., Merriam S.,
RA Lora J.M., Geddes B.C., Briskin M., DiStefano P.S., Bertin J.,
RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT activation of NF-kappa B and caspase-1-dependent cytokine
RT processing.";
RL J. Biol. Chem. 277:29874-29883(2002).
RN [2].
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lymphoma;
RA Williams K.L., Linhoff M.W., Farton J.A., Ting J.P.Y.;
RC Submitted (MAY-2002) to the ENBL/GenBank/DBJ databases.
RN [3].
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Leukocyte;
RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Spletton N., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Yoshiyuki S., Carrinci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Rahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting N., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RJ Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC activation of NF-kappa-B via IKK.
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1; Synonyms=1;
CC IsoId=P59046-1; Sequence=Displayed;
CC Name=2; Synonyms=1;
CC IsoId=P59046-2; Sequence=VSP_005524;
CC Name=3; Synonyms=1;

```

```

CC IsoId=P59046-3; Sequence=VSP_005523;
CC -!- TISSUE SPECIFICITY: Detected on y in peripheral blood leukocytes,
CC predominantly in eosinophils and granulocytes, and at lower levels
CC in monocytes.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY095146; AAM18227.1; -
CC EMBL; AY116204; AAM75142.1; -
CC EMBL; AY116205; AAM75143.1; -
CC EMBL; AY116206; AAM75144.1; -
CC EMBL; BC028069; AAM28069.1; -
CC InterPro; IPR001611; LRR.
CC InterPro; IPR007091; LRR_RNinh.
CC InterPro; IPR003590; LRR_RNinh_sub.
CC InterPro; IPR07111; NACHT_NTPase.
CC Pfam; PF00560; LRR; 2.
CC Pfam; PF02758; PAAD_DAPIN; 1.
CC PRINTS; PR00019; LEURICHRPT.
CC SMART; SM00368; LRR_R1; 11.
CC PROSITE; PS50824; DAPIN; 1.
CC PROSITE; PS50837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat; Alternative splicing.
FT DOMAIN 1 95
FT DOMAIN 211 528
FT REPEAT 713 736
FT REPEAT 742 765
FT REPEAT 770 792
FT REPEAT 799 822
FT REPEAT 827 850
FT REPEAT 884 907
FT REPEAT 941 969
FT REPEAT 998 1021
FT NP BIND 217 224
FT VARSP LIC 863 974
FT VARSP LIC 977 1032
FT CONFLICT 692 692
FT SEQUENCE 1062 AA; 120328 MW; 0AB81C87F16497F CRC64;
Query Match 25.68; Score 825; DB 1; Length 1062;
Best Local Similarity 32.64; Pred. No. 1.7e-52;
Matches 219; Conservative 104; Mismatches 245; Indels 104; Gaps 15;
QY 1 MLAGPQRLFLDGADEL-PALGQPEAAPCTDPFEASGARVLGGLSKALLPTALLVLT 59
DB 282 LIRYPERLLFIIDGFDELKPSFHDPPQPCWCLWEKRPTELLNSLRKLLPELSLT 341
QY 60 TRAAPGRLQGLCSPOCAEVRGFSKDKKKYKFPDRERRAERAYRVKVENETLFCALC 119
DB 342 TRPTALEKLHRLLEHPRHVEILGFSEAEKRYKYFHNAEQAGQVFNVDNEPLFTMC 401
QY 120 FVPVVCVICTVLRLQQLGLGRDLSRTSKTTTSVYLLFITSVLSSAPVADGPRLC--GDLR 177
DB 402 FVPLVVCVCTCQQQQLGGLRLQTSRTTIVMYLLSLMQPKFGA--PRLQPPFNQR 459
QY 178 NLCKLAREGVLGRAQFAEKLEQLERLGRSKVOTFLSKKELPGVLETVTYQFIQSSQ 237
DB 460 GLCSLAADGLWNQKILFEEQDLRKHGLDGEDV-SAFLLNMNIFOKDNCERYYSFIHLSQ 518
QY 238 EFLAALSYLLEDGGVPRTAAGVG-----TLRQDAQPHSHLVLTTRFLFGLLSAERM 290

```

519 EFFAAMYVILDEG-----EGGAGPDQDVTRLLTTEAFSPERSFLATSRFLFG:LNSETR 572
QY 291 RDIERHFGCMYSERVKQBALRWQGGCGVAPETEGAKGLEDTEEBEEBEGEPN 350
Db 573 SHLEKSLCWKYSPIKMDLLQWIGSKAQ-----SDGSTLQOGS--- 650
QY 351 YPLELLLYCLYETQEDAFVROALCRFPPELALQVRFCRMDVAVLSYCVRCPCPAGALSLIS 450
Db 611 --LFEFSLYIEGEEFIQQLSHFQIVVSNIA-SKMEHWVSFCLKRNCSAOLVRLYNG 667
QY 412 CRLVAAGKKKKS-GKRLQAS/GGGS-----SOQTTXQLPASL----- 448
Db 668 ATYSADGE-----DRARCAGAHLLVLQVRPERTVLJDAYSHLAAALCTNENLIELS 720
QY 449 -----LHPLFOAMTDPCHLSSLTLSHCKLPDAVCHDLSEALRAAPALTELGLH 498
Db 721 LYRNALGSRGKVLCCQGRHFNCKLQNLRLKRCRISSEACEDLSAALIANKYLTRMOLSG 780
QY 499 NRLSEAGRLMISEGLAWPQCRVQTVRVOLPDPQRG-LQYLVGMLRQSPALFTLIDLSSCQL 557
Db 791 NGVGFPGMWLLCEGLRHPCRLQW-LQAKCOLESAGCOEMASVLGTPTPHLVELDLTGNAL 840
QY 558 PAPVYTYLCAY-QHOGGGLQTL-----SLASVELSPQSLOELCAVKRAKPDCL 604
Db 841 EDGLRLCCQLRHPVCRRLTLKLCIKRLTAACDELASTLSVNQSLRELDLSLNEIGDL 900
QY 605 VI-----THP 609
Db 901 GVLLCEGLRHP 912

RESULT 6
CIS1_HUMAN
ID CIS1_HUMAN STANDARD; PRT; 1034 AA.
AC Q96P20; C75434; Q8TCW0; Q8TEC5; Q8WH9;
DT 28-FEB-2003 (Rel. 41, Created)
ST 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cold autoinflammatory syndrome 1 protein (Cryopyrin) (NACHT-, LRR- and
DE PYD-containing protein 3) (PYRIN-containing APAF1-like protein 1;
DE (Arginotensin)/vasopressin receptor AII/AVP-like).
GN CIAS1 OR NALP3 OR PYPAF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANTS FCAS MET-198; VAL-439
RP AND GY-627, AND VARIANT MWS VAL-352.
RX MEDLINE=21547523; PubMed=11687797;
RA Hoffman H.M., Mueller J.L., Broide D.H., Wenderer A.A.,
RA Kolodner R.D.;
RT Mutation of a new gene encoding a putative pyrin-like protein causes
RT familial cold autoinflammatory syndrome and Muckle-Wells syndrome.;
RL Nat. Genet. 29:301-305(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21909508; PubMed=11786556;
RA Marji G.A., Wang L., Geddes B.J., Brown M., Merriam S., Al-Garawi A.,
RA Mak S., Jora J.M., Briskin M., Jurman M., Cao J., DiStefano P.S.,
RA Bettin C.;
RT "PYPAF1: a PYRIN-containing APAF1-like protein that assembles with ASC
RT and activates NF-kB".
RL J. Biol. Chem. 277:11570-11575(2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), VARIANT MWS MET-198, AND
RP VARIANTS FCAS/MWS TRP-260 AND PRO-305.
RX MEDLINE=22241234; PubMed=12355493;
RA Aganna E., Martinon F., Hawkins P.N., Ross J.B., Swan J.C.,
RA Booth J.R., Lachmann H.J., Gaudet R., Woo P., Feighery C.,
RA Cotter F.B., Thome M., Hitzman G.A., Tschopp J., McDermott M.F.;
RT "Association of mutations in the NALP3/CIAS1/PYPAF1 gene with a broad
RT phenotype including recurrent fever, cold sensitivity, sensorineural

RT deafness, and AA amyloidosis".
RL Arthritis Rheum. 46:2445-2452(2002).
RN [4]
RC SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.-H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.-G., Zhang J.W.,
RA Tao J., Huang Q.-H., Zhou J., Hu G.X., Gu J., Chen S.-J., Chen Z.;
RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells".
RL Genome Res. 10:1546-1560(2000).
RN [5]
RP VARIANT FCAS MET-198, VARIANTS MWS ASN-303; MET-348; THR-439 AND
RP ARG-569, AND VARIANT FCAS/MWS TRP-260.
RX MEDLINE=21987640; PubMed=11992256;
RA Dode C., Le Du N., Cuisset L., Letourneur F., Berthelot J.-M.,
RA Vaudoir G., Meyrier A., Watts R.A., Scott D.G.I., Nicholls A.,
RA Granel B., Frances C., Garcier F., Edery P., Boulanguet S.,
RA Domergues J.-P., Delpech M., Grateau G.;
RT "New mutations of CIAS1 that are responsible for Muckle-Wells syndrome
RT and familial cold urticaria: a novel mutation underlies both
RT syndromes".
RL Am. J. Hum. Genet. 70:1498-1506(2002).
RN [6]
RP VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND
RP THR-662, AND TISSUE SPECIFICITY.
RX MEDLINE=22062556; PubMed=12032915;
RA Feldmann J., Prieur A.-M., Quartier P., Berquin P., Certain S.,
RA Cortis E., Teillac-Hamel D., Fischer A., de Saint Basile G.;
RT "Chronic infantile neurological cutaneous and articular syndrome is
RT caused by mutations in CIAS1, a gene highly expressed in
RT polymorphonuclear cells and chondrocytes".
RL Am. J. Hum. Genet. 71:198-203(2002).
CC [1]- FUNCTION: May function as a potential inducer of apoptosis.
CC Interacts selectively with apoptosis-associated specklike protein
CC containing a CARD domain (ASC). This complex may function as an
CC upstream activator of NF-kappaB signaling.
CC [2]- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=2;
CC IsoId=Q96P20-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q96P20-2; Sequence=VSP_005520, VSP_005521;
CC Name=3;
CC IsoId=Q96P20-3; Sequence=VSP_005519;
CC [3]- TISSUE SPECIFICITY: Expressed in blood leukocytes. Strongly
CC expressed in polymorphonuclear cells, undetectable or expressed
CC at a lower magnitude in B and T lymphoblasts, respectively. High
CC level of expression detected in chondrocytes. Low or no expression
CC in the other tissues tested.
CC [4]- DISEASE: Defects in CIAS1 are a cause of familial cold
CC autoinflammatory syndrome (FCAS), commonly known as familial cold
CC urticaria. FCAS is rare autosomal dominant systemic inflammatory
CC disease characterized by episodes of rash, arthralgia, fever and
CC conjunctivitis after generalized exposure to cold.
CC [5]- DISEASE: Defects in CIAS1 are a cause of Muckle-Wells syndrome
CC (MWS), a rare autosomal dominant fever syndrome with episodic
CC urticaria, arthralgia, amyloidosis and progressive sensorineural
CC deafness.
CC [6]- DISEASE: Defects in CIAS1 are the cause of chronic infantile
CC neurologic cutaneous and articular syndrome (CINCA), also known as
CC 'neonatal onset multisystem inflammatory disease', or NOKID, a
CC rare congenital inflammatory disorder characterized by a triad of
CC neonatal onset of cutaneous symptoms, chronic meningitis, and
CC joint manifestations with recurrent fever and inflammation.
CC [7]- SIMILARITY: Contains 1 DAPIN domain.
CC [8]- SIMILARITY: Contains 1 NACHT domain.
CC [9]- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC [10]- CAUTION: Ref.4 sequence differs from that shown due to frameshifts
CC in positions 893, 918 and 926.
CC -----

NAME HUMAN
 ID NAME HUMAN STANDARD; PRT: 1473 AA.
 AC Q9C000; Q9BZ28; Q9HAV8; Q9UFT4; Q9VZE0;
 DT 16-OCT-2001 (Rel. 40, Created;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE NACHT-, LRR- and PYP-containing protein 2 (Death effector filament-
 DE forming ced-4-like apoptosis protein) (Nucleotide-binding domain and
 DE caspase recruitment domain) (Caspase recruitment domain protein 7).
 GN NALP1 OR DEFCAP OR NAC OR CAFD7 OR KIA0926.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21169419; PubMed=11270363;
 RA Bertin J., Distefano P.S.;
 RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
 RT proteins";
 RL Cell Death Differ. 7:1273-1274(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21148093; PubMed=11250163;
 RA Martignon F., Hofmann K., Tschopp J.;
 RT "The pyrin domain: a possible member of the death domain-fold family
 RT implicated in apoptosis and inflammation";
 RL Curr. Biol. 11:R118-R120(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=21153743; PubMed=11078957;
 RA Hsiang T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A.,
 RA Shi M.M., Vincenz C., Ward P.A.;
 RT "Molecular cloning and characterization of DEFCAP-3 and -S, two
 RT isoforms of a novel member of the mammalian Ced-4 family of apoptosis
 RT proteins";
 RL J. Biol. Chem. 276:9230-9238(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACT-CN.
 RX TISSUE=Testis;
 RX MEDLINE=21153744; PubMed=1113115;
 RA Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewska M., Krajewski S.,
 RA Godzik A., Reed J.C.;
 RT "A novel enhancer of the Apaf1 apoptosome involved in cytochrome
 RT c-dependent caspase activation and apoptosis";
 RL J. Biol. Chem. 276:9239-9245(2001).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX TISSUE=Brain;
 RX MEDLINE=99246063; PubMed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hiroseawa M.,
 RA Miyajima N., Tanaka A., Kohani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XI--
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro";
 RL DNA Res. 6:63-70(1999).
 RN [6]
 RP SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).
 RC TISSUE=Uterus;
 RC Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Able to form cytoplasmic structures termed death
 CC effector filaments. Enhances APAF1 and cytochrome c-dependent
 CC activation of pro-caspase-9 and consecutive apoptosis. Seems to
 CC bind ATP.
 CC -!- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9
 CC and with APAF1 in a cytochrome c-inducible way leading to the
 CC formation of an apoptosome. This interaction may be ATP-dependent.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=1; Synonyms=NAC beta, DEFCAP-L;

ISOId=Q9C000-1; Sequence=Displayed;
 Name=2; Synonyms=NAC alpha, DEFCAP-S;
 ISOId=Q9C000-2; Sequence=VSP_004327;
 Name=3; Synonyms=NAC gamma;
 ISOId=Q9C000-3; Sequence=VSP_004326;
 Name=4; Synonyms=NAC delta;
 ISOId=Q9C000-4; Sequence=VSP_004326;
 TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are
 expressed in peripheral blood leukocytes, chronic myelogenous
 leukemia cell line K-562, followed by thymus, spleen and heart.
 Also detected in lung, placenta, small intestine, colon, kidney,
 liver and muscle.
 -!- SIMILARITY: Contains 1 DAPIN domain.
 -!- SIMILARITY: Contains 1 NACHT domain.
 -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@sib-sib.ch).
 EXBL; AF298548; AKG15254.1; -
 EXBL; AF310105; AAG30288.1; -
 EMBL; AF229059; AAK00748.1; -
 EMBL; AF229060; AAK00749.1; -
 EMBL; AF229061; AAK00750.1; -
 EMBL; AF229062; AAK00751.1; -
 EMBL; AS023143; BAA76770.1; -
 EMBL; AL111470; CAB55945.1; -
 PIR; T17255; T17255.
 DR HSSP; P13489; I44Y.
 DR KIM; G05636; -
 DR GO; GO:0005622; C:intracellular; IC.
 DR GO; GO:0016506; F:apoptosis activator activity; NAS.
 DR GO; GO:0008656; F:caspase activator activity; NAS.
 DR GO; GO:0019899; F:enzyme binding activity; IPI.
 DR GO; GO:0006919; P:caspase activation; NAS.
 DR GO; GO:0006917; P:induction of apoptosis; NAS.
 DR InterPro; IPR001315; CARD.
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR_RNinh.
 DR InterPro; IPR007091; NACHT_NTPase.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR_2.
 DR Pfam; PF02758; PAAD_DAPIN_1.
 DR PRINTS; PR00364; DISEASERSIST.
 DR PROSITE; PS0209; CARD; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 DR PROSITE; PS0837; NACHT; 1.
 KW Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 KW Alternative splicing.
 FT DOMAIN 1 92 DAPIN.
 FT DOMAIN 328 637 NACHT.
 FT REPEAT 704 725 LRR 1.
 FT REPEAT 807 830 LRR 2.
 FT REPEAT 864 887 LRR 3.
 FT REPEAT 921 944 LRR 4.
 FT REPEAT 950 973 LRR 5.
 FT REPEAT 1199 1215 LRR 6.
 FT REPEAT 1216 1236 LRR 7.
 FT DOMAIN 1374 1463 CARD.
 FT NP_BIND 334 341 ATP (POTENTIAL).
 FT VARSP_LIC 958 987 Missing (in isoform 3 and isoform 4).
 FT VARSP_LIC 1262 1305 /FTId=VSP_004326.
 FT VARSP_LIC 340 340 Missing (in isoform 2 and isoform 3).
 FT MUTAGEN 340 340 K->L: NO EFFECT.
 FT MUTAGEN 340 340 K->S: NO EFFECT.

```

FT CONFLICT 155 155 I -> H (IN REF. 1).
FT CONFLICT 246 246 T -> S (IN REF. 1).
FT CONFLICT 782 782 T -> S (IN REF. 1).
FT CONFLICT 878 878 T -> M (IN REF. 1).
FT CONFLICT 995 995 T -> I (IN REF. 1).
FT CONFLICT 1119 1119 M -> V (IN REF. 1).
FT CONFLICT 1184 1184 M -> V (IN REF. 1 AND 6).
FT CONFLICT 1241 1241 V -> L (IN REF. 1).
FT CONFLICT 1366 1366 E -> C (IN REF. 1).
SQ SEQUENCE 1473 AA; 165865 MW; 438F0DCB45C2562D CRC64;

Query Match 23.7%; Score 762; DB 1; Length 1473;
Best Local Similarity 33.2%; Pred. No. 1e-47;
Matches 220; Conservative 103; Mismatches 207; Indels 132; Gaps 22;

QY 1 MLAGPRLFLILDGADLPP--ALGGPEAAFPCTDPFEAASGARVLGGLLSKALLPTALLV 58
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 ILSPERLLFILDGVDE-PGWVLEPSSSELHWSQPQADALLGSLGKT-LPEASFLI 456
QY 59 TTEAAPGRLOGLCSPOCAEVRGFSDDKKYFYKFFRDERAERAYRVKENETLAL 118
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 TARTTALQNLPSLEQARWVEVLGFSSESKKEYFYRYFTDERQAIRAPFLVSKNELWAL 516
QY 119 CFVPFVCMIVCTVLRQOLELGRDLSRTSKTTSVLLFTSVLSAPVADGPRLOGDLRN 178
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 CLVPWWSKLACTCLMQMKKEKLTSTKTTTLCHLYLAQALQAPL--GPO---LRD 570
QY 179 LCLAREGVLRGAQFAEKELEQLRGSVKVQTLFLSKKELPGVLETE---VYQFIDQS 235
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 LCSAAEGWIKKTTLSPDLRKHG-DGAIIST-FLKX---GILQEHIPLSYFIHL 625
QY 236 FOEFLAALSLLLEDGGVPRTPAAGGVCTLRGDAQPHSHLVL-----T 277
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 PQEFPFAXSVLEL-----EKGRG-----KHSNCIIILEKTLKXGTHG-LFGAS 670
QY 278 TRFLFGLLSAERYRDIERHFGCMVSRVQEAALRWVQGGQCPGVAPEVTEGAKGLETT 337
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 TRFLGLLSDEGEREMENIFHCRLSQ--GRNLQWV-----ESLO-L 710
QY 338 EEPDEEEEGEENPYPLEL-YCLVETOEDAFVRQALCFEPBELALQVRFC---RYCVAVLS 394
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 LQP-----HSLSEHLCLVETRNKTLFTLOVMAHFEBMGX-----CVETDMLLVCT 755
QY 395 YVVRCCPAGALRLISCR-----LVAAQPKKKKSLGKRLQ 429
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 FCIKFSRHVKKLQIEGRCHRSWTSPTMVV-FRWVPVTDAYNCILPSVLKV-RNL-KELD 814
QY 430 ASLGGSSQGTTKQLPASLLHPFQAMTDLCHLSLTLSHCKLPDAVCRDLSELRAP 489
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 LS-----LHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQ 865
QY 490 ALTELGLLHNLSEAGLRMLSEGLAWPQCRVQTVRVOLPFGRL-----OYLGMRLQS 544
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 TLELDLSPNVLDVAGAKHLCRLRQPSKLO--RLQL--VSCGLTSDCCQQLASVLSAS 921
QY 545 PALTTLDLSCQCPAPMVTVYLZAVLHQCGGLQTLASVELSEQLQELQVAKRAKPD 604
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 PSLKELDLQNNLDVGVRLTLEGLRHPACKLRILGLDQTLTSLDEMQRLEALEGEKPOL 981
QY 605 VI 606
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 982 LI 983

RESULT 8
PYA3_HUMAN STANDARD; PRT; 980 AA.
AC Q8WX94;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE PYRIN-containing ARAFI-like protein 3.
GN PYPAF3.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22182427; PubMed=12019269;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA Lora J.M., Geddes B.J., Brisken M., DiStefano P.S., Bertin J.;
RT "PYPAF3, a novel PYRIN-containing ARAFI-like protein that regulates
RT activation of NF-kappa B and caspase-1-dependent cytokine
RT processing.";
RT J. Biol. Chem. 277:29874-29880(2002).
AL J. Biol. Chem. 277:29874-29880(2002).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
CC ENBL: AF464765; AAL69963.1; --
CC InterPro: IPR007091; LRR_RNinh.
CC DR InterPro: IPR007111; NACHT_NTPase.
CC DR InterPro: IPR004020; PAAD_DAPIN_dom.
CC DR Pfam: PF02758; PAAD_DAPIN; 1.
CC DR PROSITE: PS50824; DAPIN; 1.
CC DR PROSITE: PS50837; NACHT; 1.
CC KW ATP-binding; Leucine-rich repeat; Repeat.
CC FT DOMAIN 1 93 DAPIN.
CC FT DOMAIN 172 491 NACHT.
CC FT REPEAT 614 638 LRR 1.
CC FT REPEAT 674 697 LRR 2.
CC FT REPEAT 760 784 LRR 3.
CC FT REPEAT 788 810 LRR 4.
CC FT REPEAT 817 840 LRR 5.
CC FT REPEAT 845 868 LRR 6.
CC FT REPEAT 874 897 LRR 7.
CC FT REPEAT 902 928 LRR 8.
CC FT REPEAT 933 957 LRR 9.
CC FT NP_BIND 178 185 ATP (POTENTIAL).
CC SQ SEQUENCE 980 AA; 111806 MW; 822AF2FD4338003D CRC64;

Query Match 20.2%; Score 650; DB 1; Length 980;
Best Local Similarity 28.6%; Pred. No. 9.6e-40;
Matches 221; Conservative 102; Mismatches 256; Indels 194; Gaps 23;

QY 1 MLAGPRLFLILDGADLPPALGGPEAAFPCTDPFEAASGARV-LGGLLSKALLPTALLV 59
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 ILAQQRILFVVDGLDELKVPFGALLQIDICGWKKKPPVPLGLSKLKKMLPRALLVT 300
QY 60 TRAAPGRLOGLCSPOCAEVRGFSDDKKYFYKFFRDERAERAYRVKENETLALC 119
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 TRPRALRDLQLLAQQPIYVRVEGLEEDRRAYFLRHFGDEDOAMRAFELMRNAALFOLG 360
QY 120 FVPFVCMIVCTVLRQOLELGRDLSRTSKTTSVLLFTSVLSAPVADGPRLOGDLRN 179
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 SAPAVCMIVCTVTKLQWKEGEDPVPTCLTRTGLRFLCS-----RFPQAGLRGALRTL 415
QY 180 CRLAREGVLRGAQFAEKELEQLRGSVKVQTLFLSKKELPGVLETEVYQFIDQSOF 239
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 SLIAAQGLNAQMSVFRREDLERLGVQESDLR-LFLDGDILRQDRVSKGCGYSIHLSFQOF 474
QY 240 LAALSYLE-DGGVPRTPAAGGVCTLRGDAQPHSHLVLTTRFLGLLSAERMROI 293
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D5 LTALFVALEKEGEDRDGHAWDIDGVQKLLSGEERLKNPDLIQVGHFLGLANERAKEL 534
QY 294 ERHFGCMVSRVQEAALRWVQGGQCPGVAPEVTEGAKGLEDTPEPEEEEGEENPYL 353
D5 :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


Db 535 EATGCRMSPIKQELLO-----C-----KALHANKP 563
Cy 354 -----ELLYCYETOEDAFYQALCRPELAL-----CRVFCR-----MDVAIVLSYC 396
Dz 564 SVTDLKEVLGCLYESQEEELAKVAVAPPEKIS:HLTKTSEVYHCSFSLKHCOLOKLSLO 623
Qy 397 V-----KC-----CP--AGQALRLIS-----CRLVAAQKKK----- 421
Db 624 VAKGVFLENYMDFE:CIFFERCTYLTIPNWARODLSLELMTDFCSLFSNSNKLFLVKK 683
Cy 422 -----KSLGK3--LQASLGG--- 434
Db 634 QSLSDSSVRLCHVTRSTCHLQKVEIKVTPDTAYRDFCLAFIGKKTLLTHLAGHIE 743
Qy 435 -----GSSQGTQK-----LPASLLHP 451
Db 744 WERTWMLMLCOLLRNHNKCNQYLRGLGHGHCATPEQMAEFYVYKANKQSLKHLRSANVLD 803
Qy 452 -----LFOAMTDPCLHSLSLTSLHCHKLPDANCRLDSEALRAAPA:TELGLLHNLSEAGL 506
Db 804 EGAMLSYKMTTRPKFIFLQWLSLGNCRLETSACKDLAAVTVSKKTLHLCLAKNPIGTGV 863
Qy 507 RYLSSEGLAMPQCRVOT-VRVQDPDPQRLQYVGMCRQSPALTLID:SGCQCPAPMYTY 565
Db 864 KFLCEGLSVDPCKQLTLV:QCSITKLCGRVYSEALQEA:CSLTNLDLSINQI-ARGLWIL 922
Qy 566 CAVLHQHCGGLQTLASVEISEQSLQELQAVKRAKPOLVITHPALDGHPOPP 619
Db 923 COALENPNCKLHKLKTYETN:EIKKLLEEVKKNPKLTIDCNA-SCATAPP 974
RESULT 9
NAL2_HUMAN STANDARD: PRT; 1062 AA.
ID NAL2_HUMAN STANDARD: PRT; 1062 AA.
AC Q9NWX02; Q9BWN5; Q9H6G6; Q9HAV9; Q9NWK3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NACHT-, -RR- and PYD-containing protein 2 (Nucleotide-binding site
DE protein 1) (PYRIN-containing APAF1-like protein 2).
GN NALF2 OR NES1 CR PYPAF2.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21169419; PubMed=11270363;
RA Bertin J., Distefano P.S.;
RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
RL proteins.";
RL Cell Death Differ. 7:1273-1274(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=21148093; PubMed=11250163;
RA Martinon F., Hofmann K., Tschoopp J.;
RT "The pyrin domain: a possible member of the death domain-fold family
RT implicated in apoptosis and inflammation.";
RL Curr. Biol. 11:R118-R120(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22162427; PubMed=12019269;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Verriam S.,
RA Lora J.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.;
RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT activation of NF-kappa B and caspase-1-dependent cytokine
RT processing.";
RJ J. Biol. Chem. 277:29874-29880(2002).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Colon, and Kidney epithelium;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,

RA Nakamura Y., Isegai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung, and Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.C.,
RA Hopkins S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schettz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Sosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RJ Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May be implicated in apoptosis (By similarity).
CC -!- COFACTOR: Binds ATP (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC isoId=Q9NWX02-1; Sequence=Displayed;
CC Name=2;
CC isoId=Q9NWX02-2; Sequence=VSP_005522;
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF298547; AAG15253.1; ALT_INIT.
DR EMBL; AF310106; AAG30289.1; -.
DR EMBL; AF464764; AAL69962.1; -.
DR EMBL; AK000517; BAA91223.1; -.
DR EMBL; AK025952; BAB15293.1; -.
DR EMBL; AK003784; BAA91377.1; ALT_INIT.
DR EMBL; BC003592; AAH03592.1; -.
DR EMBL; BC001039; AAH01039.1; -.
DR HSSP; P10775; 2BNH.
DR InterPro; IPR007091; LRR RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
KW Alternative splicing.
FT DOMAIN 1 94 DAPIN.
FT DOMAIN 207 526 NACHT.
FT REPEAT 467 491 LRR 1.
FT REPEAT 622 645 LRR 2.
FT REPEAT 754 777 LRR 3.
FT REPEAT 810 832 LRR 4.
FT REPEAT 839 862 LRR 5.
FT REPEAT 867 890 LRR 6.
FT REPEAT 924 947 LRR 7.

```
FT REPEAT 981 1005 LRR 8.
FT REPEAT 1010 1033 LRR 9.
FT NP_BIND 213 220 ATP (POTENTIAL).
FT DOMAIN 518 523 POLY-GLU.
FT VARSPIC 133 154 Missing (in isoform 2).
FT CONFICT 1 1 N -> V (IN REF. 1).
FT CONFICT 35 35 L -> P (IN REF. 1).
FT CONFICT 304 304 I -> S (IN REF. 4: BAB15293).
FT CONFICT 364 364 E -> K (IN REF. 1).
FT CONFICT 980 980 MISSING (IN REF. 1).
FT CONFICT 1052 1052 NISSING (IN REF. 5).
SQ SEQUENCE 1062 AA; 120514 MW; 4DBB0F629C2BC8A7 CRC64;

Query Match 19.2%; Score 616.5; DB 1; Length 1062;
Best Local Similarity 27.3%; Pred. No. 3e-37;
Matches 204; Conservative 104; Mismatches 260; Indels 179; Gaps 18:

QY 1 MLAQORLLFLDGADELPAALGGPEAPCTDPFEAASGARV-LGGILLSKALLFTALELVT 59
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
276 ILAQARKLFLVIDGFDLGA-PGALLIEDICGDEKXKPVVLLGSLNVRVMLPKAAJLV 335
QY 60 TRAAAPRIGRGLCSPOCAEVRPFSKDKKKYKFFRDERAERAYRVKENTLFALC 119
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
336 TRPRALDLRLAEPIYIRVEGFLEEDRRAYFLRHFGDEQAMRAFELMRNAAJFCIG 395
QY 120 FVPFVCKIVCTLRQOQLEGRD-SRTSKTTTSVYLLFITSVLSSAPVADGPRLOGDLRL 179
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
396 SAPAVCKIVCTLRQOQLEGRD-SRTSKTTTSVYLLFITSVLSSAPVADGPRLOGDLRL 450
QY 180 CRLAREGLVGRRAQFAEKLEQLELRGSKVOTLFLSKELPGVLETVTYQIFDOSFOF 239
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
451 SLLAQGLWAQTSVLHREDLER-LGVQESDLR-LFLQDGLLRQDRVSKGCVSFHLSFQOF 509
QY 240 LAALSYLE-----DGGVTRTAAGCVTLTAG-DAQPHSHLVLTTRFLGLLSAERNR 291
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
510 LTALFYTLKEEEDRDGHT--WDIGDVQKLSSGVERLRNPDLIQAGYYSFGLANEKRAK 567
QY 292 DIERHFGCVSERVKOEALRWVQGGCGFGVAPEVTEGAKGLEDTEPEEEEGEPEEPY 351
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
568 ELEATFGCRMSPIKOELELC-----DISCKGHSTVTD-----LQ----- 603
QY 352 PUELLVCLYTOEDAFVQALCRFPPEALQVRFCRMDVAVLSYVRCPPAQALRL--- 408
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
604 --ELLCGLVESQEEELVKEMWAQFKEISL--HLNAVDDVVPSSFCVKHCRNLQKXSLQVI 658
QY 409 -----I 409
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
659 KENLPENVTASESDAEVRSQDQHMPLPFWTDLCSIFGSKNKLXGLAINDSFLSASVRI 718
QY 410 SCLRLVAA-----CEKKKSLIG-----KRLQASLGGGSS-----QCTTK--QLPA---SL 448
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
719 LCEQIASDCHLQVVVFKMISPADARNLCLALRGHKVTYV-LQGNDDQDQFPAJCEVL 778
QY 449 LAP-----LFOAYTDP 459
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
779 RHPECNRLVGLVSCSATTQWADLSALEVNGSLTCVNLSDNELLDEGAKLYTLAR? 838
QY 460 LCHLSSETLSHCKLPDAVCEDLSAARAPALTEGLLENRLSEAKRLMSGLKAPQCR 519
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
839 KCFLOQLSENCHLTETANKDAAVLVVSRETHCLAKNPIGNTVKVFCEGLGRYPCK 898
QY 520 VQTVRVCLPD-PORGILQYLVGLMRQSPALTTLDLSCCOLPAPVTVYLCVAVHQSGCGL? 578
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
899 LQTLVLWNCDDITSDGGCDLTKLQEKSSLLCLDLGLNHLGVKMKFCEALRKP-CNLR 958
QY 579 LSLASVELSEQSLQELQAVKRPDVL 605
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
959 LNLWGCISPPFSCDCLCSALSCQSIV 985
```

RESULT 10

NAL4_HUMAN

```
ID ID NAL4_HUMAN STANDARD; PRT; 994 AA.
AC Q96M22; O96AY6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE NACHT, LRR- and PYD-containing protein 4 (PAA4 and NACHT-containing protein 2) (PYRIN-containing APAF1-like protein 4) (Ribonuclease inhibitor 2)
GN NALP4 OR PYPAF4 OR PAN2 OR RNH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Martinson F., Tschoop J.;
RT "NALP4 a novel member of the PYD, NACHT, and LRR family.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Fiorentino L., Reed J.C.;
RT "Pan2, a novel PAA4-containing protein.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Bertin J.;
RT "PYPAF4: a novel PYRIN-containing APAF1-like protein.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Miyamoto T.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H., Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K., Kurehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RT "NED0 human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 436-994 FROM N.A. (ISOFORM 2).
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Aitschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Haieh F., Diatchenko J., Narusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Lquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalob D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalusz D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Comment=Experimental confirmation may be lacking for some isoforms;
CC Name=1;
```



```
-----
EMBL: AY054986; AAL15549.1; --
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh..
DR InterPro; IPR003593; LRR_RNinh_sub.
DR InterPro; IPR007111; NACHT_NTFase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF003560; LRR_3.
DR Pfam; PF02258; PAAD_DAPIN; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00368; LRR_R1; 11.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 57 148 DAPIN.
FT REPEAT 280 602 NACHT.
FT REPEAT 704 727 LRR 1.
FT REPEAT 730 753 LRR 2.
FT REPEAT 780 803 LRR 3.
FT REPEAT 809 832 LRR 4.
FT REPEAT 836 863 LRR 5.
FT REPEAT 865 892 LRR 6.
FT REPEAT 893 916 LRR 7.
FT REPEAT 950 973 LRR 8.
FT REPEAT 979 1002 LRR 9.
FT REPEAT 1007 1034 LRR 10.
FT REPEAT 1036 1059 LRR 11.
FT REPEAT 1064 1092 LRR 12.
FT REPEAT 1121 1142 LRR 13.
FT NP_BIND 286 293 ATP (POTENTIAL).
SQ SEQUENCE 1200 AA; 134235 MW; 9A070D2A771B28FA CRC64;

Query Match 15.8%; Score 507.5; DB 1; Length 1200;
Best Local Similarity 26.7%; Fred.No. 3.1e-299; Indels 117; Gaps 14;
Matches 187; Conservative 98; Mismatches 239;

QY 1 MAAQORLLFLDGADELPAIGPEAPCTDPPEAASGARVLGGLSKALLFTALVTT 63
DB 350 INSRPERLLFIIDGFDDGLSVLNDT;KLCQDAEKQPPFTLRSLRLKVLFPESFLIVT 409
QY 61 RAAAPGRLGRLCSPOCAEVRGSDKKKFKYKPRDRRAEAVRFVKNET;PALCF 123
DB 420 RYVGTGLKSEVVSPPYLLVFGISGQRTHLLERGI;GEHQKTCGLRAIINKRELLDQCO 469
QY 121 VPFCVIVCTVLRQQLGRLGRLSRTSKTTTSVLLFTITSVLSPAYAD---GPRQGDLR 177
DB 470 VPAVGSLLVALQQLQDVUESVAPFQNTLTGLAAAFHQL;PRGVVRCLNLSERVJK 529
QY 178 NLGRJAREGVLRRAQFAEKELEQLRLGSKVQTLFLSKKELPGV;LETETVYCFIDQSPQ 237
DB 530 RFCRMAVEGVWNRKSVFDGDL;MVQGLGESELRALFHMNILLPD-SHCSEYVTFPHLSQ 598
QY 238 EFLNALSYLLEDGCV-----FRTAAGVGVTLLRGDAQPHSHLVLTTRFLGILLSAERNR 292
DB 589 DFCAALYVYVLEGLIEIPALCFVYVTKRSMELKQAGFHHSMMKRF;FGLVISEDVRP 648
QY 293 IERHFGGMWSERVQKQALRWVQGGCGCPVAPEVTEGAKGLEDTPEPEEEGEPNPV 352
DB 649 LEVLLGCPVPLGVKQLLHVLSLQ-----QPNAT-----TPGDT 684
QY 353 LELLYCLYETQEDAFVQALCRPELALQVRFCRMDVAVLSYVVRCCPAGQALRL----- 408
DB 685 LDAFHCLFETQDKREFVRLALNSFOEVLPIQ--NLDLIASSFC;CHCPYLRKIRVDVKG 742
QY 409 -----ISCLVAQEKKKSLGKRLQ-----ASLGGSS--QGTTKQLP 445
DB 743 IFRPDSAEACPVVPLMMRDKTLIEQOWEDFCSMLGTHPHLRQLDLGSSILTERAMKTL 802
QY 446 ASLLHP-----LF-----QAVT 457
DB 803 AKLRHPTCKTLQTLNFRNAQITPGVQH;NRIVMANRNLSNLGGTHLKEEDVRMACEALK 862
QY 458 DPLCHLSLTLSHCKLPDAVCRDLSEALRAAPALTEGLLHNRLSEAGLRM-SEGILAMPQ 517
```

```
DB 863 HPKCLLESRLDCGGLTHACYLVKISQILTTSPSLKSLAGNKVTDGCVTPLSDALEVSQ 922
QY 518 CRVQTRVQLPD---PQRGLQYLVGMRLQSPALTTDLSCQCLPAPMYTYLCAVLQHGQC 574
DB 923 CALQ--KLILEDGGITATGCSLASALVSNRSLTHLCLSNNSLNEGVLNLCRSMRLPHC 980
QY 575 GLOTLSLASVELSEQSLQELQAVKRAKPDVITHPALDGH 615
DB 981 SLQRLMLNQCHLDTAGCGSLALALMGNSWL--THLSLSMNP 1019

RESULT 12
MATE_MOUSE
ID MATE_MOUSE STANDARD; PRT; 1111 AA.
AC OPRIMS; Q9JLR2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Maternal antigen that embryos require (Mater protein) (Ooplasm-
DE specific protein 1) (OPL).
GN MATER.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH Swiss; TISSUE=Ovary;
RX MEDLINE=99360614; PubMed=10433232;
RA Tong Z.-B., Nelson L.M.;
RA "A mouse gene encoding an oocyte antigen associated with autoimmune
RA premature ovarian failure.";
RX Endocrinology 140:3720-3726(1999).
RN [2]
RP SEQUENCE FROM N.A.; AND VARIANTS ASP-7; LEU-42; LYS-87; PRO-256;
RX GLN-977 AND 1003-ASN-ASN-1004.
RC STRAIN=129/Sv;
RX MEDLINE=20222768; PubMed=10754103;
RA Tong Z.-B., Nelson L.M., Dean J.;
RA "Mater encodes a maternal protein in mice with a leucine-rich repeat
RA domain homologous to porcine ribonuclease inhibitor.";
RL Mamm. Genome 11:281-287(2000).
RN [3]
RP FUNCTION.
RX MEDLINE=20517328; PubMed=1062459;
RA Tong Z.-B., Gold L., Pfeifer K.E., Dorward H., Lee E., Bondy C.A.,
RA Dean J., Nelson L.M.;
RA "Mater, a maternal effect gene required for early embryonic
RA development in mice.";
RL Nat. Genet. 26:267-268(2000).
CC -!- FUNCTION: Necessary for embryonic development beyond the 2-cell
CC stage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Oocyte-specific.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF074018; AAD51762.1; -.
DR EMBL; AF143573; AAF64393.1; -.
DR EMBL; AF143559; AAF64393.1; JOINED.
DR EMBL; AF143560; AAF64393.1; JOINED.
DR EMBL; AF143561; AAF64393.1; JOINED.
DR EMBL; AF143562; AAF64393.1; JOINED.
DR EMBL; AF143563; AAF64393.1; JOINED.
```



```

RX MEDLINE=22198257; PubMed=12477932;
RA Strausberg R.J., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Induces NF-kappaB via RICK (CARDIAX, RIP2) and IKK-
gamma. Confers responsiveness to intracellular bacteria;
CC lipopolysaccharides (LPS) (By similarity).
CC -!- SUBUNIT: Binds to RICK by CARD-CARD interaction (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event:Alternative splicing; Named isoforms=2;
CC Name=1;
CC Name=2;
CC Name=3;
CC IsoId=Q8K3Z0-1; Sequence=Displayed;
CC IsoId=Q8K3Z0-2; Sequence=VSP_007069, VSP_007070;
CC Note:No experimental confirmation available;
CC -!- SIMILARITY: Contains 2 CARD domains.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 10 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF520774; AA076073.1;
CC EMBL: BC044774; AA044774.1; ALT_INIT.
CC MGD: MGI:2429397; Card15.
CC InterPro: IPR001315; CARD.
CC InterPro: IPR007091; LRR_RNinh.
CC InterPro: IPR007111; NACHT_NTPase.
CC Pfam: PF00560; LRR_2.
CC PROSITE: PS50209; CARD; 2.
CC PROSITE: PS50837; NACHT; 1.
CC Alternative splicing;
CC DOMAIN 6 104 CARD 1.
CC DOMAIN 16 200 CARD 2.
CC DOMAIN 223 600 NACHT.
CC REPEAT 685 709 LRR 1.
CC REPEAT 726 749 LRR 2.
CC REPEAT 766 792 LRR 3.
CC REPEAT 794 817 LRR 4.
CC REPEAT 822 845 LRR 5.
CC REPEAT 850 873 LRR 6.
CC REPEAT 906 929 LRR 7.
CC REPEAT 934 962 LRR 8.
CC REPEAT 963 985 LRR 9.
CC REPEAT 1005 1019 LRR 10.
CC VARSPLIC 1 7 Missing (in isoform 2).
CC VARSPLIC 195 195 /FTId=VSP_007069.
CC E -> EGYSLCKNSRCDRGFT;CLPCL (in isoform
2).

```

```

FT VARIANT 212 212 /FTId=VSP_007070.
FT VARIANT 240 240 T -> A (in strain NMRI).
FT VARIANT 422 422 Q -> R (in strain NMRI).
FT VARIANT 422 422 L -> C (in strain NMRI).
FT VARIANT 485 485 G -> V (in strain NMRI).
FT VARIANT 603 603 V -> A (in strain NMRI).
FT VARIANT 675 675 V -> I (in strain NMRI).
FT VARIANT 925 925 E -> Q (in strain NMRI).
SQ SEQUENCE 1020 AA; 113561 MW; 25504905ECF70FBB CRC64;

Query Match 9.5%; Score 305; DB 1; Length 1020;
Best Local Similarity 23.5%; Pred. No. 1.6e-14;
Matches 157; Conservative 111; Mismatches 277; Indels 124; Gaps 28;

QY 1 MLAQPORLLFILDGDELPAALGPEAACTDFEASGARVLGGLLSKALLFTALLVTT 60
DB 347 LLDHPORVLLTFDGLDEFKFRFTDRRHCS-PIDPTSVQTLLENLQGLNKKACKVLTS 405
QY 61 RAAAPGRLOGRLCSPQCAEVRGFSDDKKKFKYKFFRDERRAEAYRVFKENETLPAICF 120
DB 406 RPDAYSALLRKFKVTE-LQKGFSEGIQLYLRKHREPGVADRLLIQATSLHGLCF 464
QY 121 VPFVWIVCTVLRQOLELGRDLRSKTTTSVYLLPITSVL-----SSAPVADGRL-- 172
DB 465 LPVFSMMVSRCHRELLQNRFPTTS--TMYLLILOHFLHASPPOSSPLGLGGLIQ 521
QY 173 --QGLRNLCRLAREGVLCGRRAQFAEKELEQLERGSKVOTLFL--SKKELPGV-ETEV 228
DB 522 SRLSTLLHGLHLAGLAMSVCVFSQAQQAQVDDDISLGLFLVRAQSSVFG--SKAP 578
QY 229 YGFIDQSFQEFLLA-----LSYLLLEDGGV-----PRTAAGG 259
DB 579 LFLHTTFOCFFAFVLAVSADTSVASLKHFLSCGLGSSLLGRLLPNLCIGSRVKKGS 638
QY 260 VGTLLRGDAQPHSHLVLTTRFLGLLSAERMMDIERHFGCMVSRVKOEALRWQGGGG 319
DB 639 EAALLQ-KAEPH-NLQIT-AAFLAGLLS-QCHRDLLA--ACQVSEV-----LLQROARARS 689
QY 320 CPQVAPVETGAKGLDTEPEEEBEGEPNYPLELYCYLVEODAFVQALCRFFELA 379
DB 690 C--LAHSLREHFSIPPVPGGETKSHAMPGF-IWLIRSLYENQEEQLAQEAVERL-DIG 745
QY 380 LQVRPFCRM--DVAVLSYCVRCCPAGQALRLISCLVAAQEKKKKSLG-----KRLQASL 432
DB 746 HLKLTCTCRVGAECALAFVLQHLQRPVALQL-----DYSVGDVGVEQLRPCL 794
QY 433 G-----GGSSQGTTKQPLASLLHPLFOAMTQPLC-HLSSLTLSHCKLPDVCVRL 482
DB 795 GVCTALYLRDNNISDRGARTLVECAL-----RCEQLQKLALFNKLTDAACSM 843
QY 482 SEALRAAPALTELGLHNLSEAGLRMLSEGLA-----WPCRVQTVRVOLPDFQ 532
DB 844 AKLLAHKQNFSLRVGNHHTAAGAEVLQAQGLKSNSTLSKFLGFWGN-----SVGD-- 893
QY 532 RGLQYLVGMLRSPALTTLLDLSCGLPAPMVTYLCVQLQCGGLQTLASVSEBSQL 591
DB 894 KGTQALAEVVDHQNKLMLSLVGNNI-GSGAGAEALALMLKXNLSLEELCLEENHICDEG 952
QY 592 QEL-QAVKR 599
DB 953 YSLAEGLKR 961

RESULT 15
C2TA_HUMAN STANDARD; PRT: 1130 AA.
ID C2TA_HUMAN
AC P33076;
CT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE MHC class II transactivator (CIITA).
GN MHC2TA OR CIITA.
OS Homo sapiens (Human).

```


QY 609 PALDGHPPPK 619
DB 472 SKLGRHPPPK 492

RESULT 2

TI7255
hypothetical protein DKFZp586O1822.1 - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
A:Accession: TI7255
R:Kocher, K.; Beyer, A.; Xaves, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: TI7255
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-192 <KOE>
A:Cross-references: EMBL:AL17470
A:Experimental source: adult uterus; clone DKFZp586O1822
C:Genetics:
A:Note: DKFZp586O1822.1

Query Match 23.7%; Score 762; DB 2; Length 1192;
Best Local Similarity 33.2%; Pred. No. 1.1e-49;
Matches 220; Conservative 103; Mismatches 207; Indels 132; Gaps 22;

QY 1 MLAQPORLFILOGADELP--ALGQPEAAPCTDPFEAASGARVLGGLLSKALLPTALLV 58
DB 117 ILSRPERLFILOGVDE-PGVVLOEPSELCHWSQOPADALGSLGKTLFEASF 175
QY 59 TTRAAAPRLQGRLCSPCAEVRGFSKDKKKYFKFPRDERRAERAYFVKENETLPAL 118
DB 176 TARTALONLPSLEQARWVULGVSESSRKEYRYFTDERQAIRAFRLVKSKEWAL 235
QY 119 CFVPVWVICTVLRQQLGELGRDLSRTSKTTTSVYLLFTSVLSAPVADGPRLOGSLRN 178
DB 236 CLVPVWSLACTCLMQQMKREKLTSTKTTLCHVLAQALCAQPL--GFC---LRD 289
QY 179 LCRLAREGVJRRAGFAKELEQLERKSKVQTLFLSKKELPGVLETE---VTQFDDQS 235
DB 290 LCSLAAREGIMOKTLFSPDCRKHSJGDAIIST-FLGM---GILQHPILPSYSF-HLC 344
QY 236 PQEFALALSYLLEDCGVPRTAAGGVGTLRGDQAPSHLVL-----T 277
DB 345 PQEFAMSYVLED-----EKGRG-----KHSNCIIDLEKTEAYGIHGLFAS 369
QY 278 TRFLPGLLSAERMRIERHFGCMVSRVKQBALRWVQGGCGGCGVAPVETEGAKGLEDT 337
DB 390 TRFLGLLSDEGERENENIFHRLSQ--GRNLMQWV-----PSLQLL 429
QY 338 EEPSEEEEGEENYPLELLCYLCTOEDAFVROALCRPELALORVAFV---RYDVAVL 394
DB 430 LQP-----HSLSLRCLYETRNKTLFTQVNAHFEEMG---CVETVVELLVCT 474
QY 395 YCVRCPPAGQALRLISCR-----LVAAQKKKSLGKRLQ 429
DB 475 PCIKFSRHVKKQLLEGQRHSTNSPTMVLFVRWVPVTDAYWQILFVLKTRNL-KELD 533
QY 430 ASGGGSGGTGKQLPASLLHPLFOAMTDPLCHSLSLTSLSHCKLPDVCRLDSEALRAAP 489
DB 534 LS-----LHSAVKSCLKTLRRPRCLLETLRAGCGLTAEDCKDLAFGLRANQ 584
QY 490 ALTELGLLHNLSEAGRLMSEGLAWPCRVQTVRVOLPDQRLG-----QYLVGKLRQS 544
DB 585 TULELGLSFGNVLTDAGAKHLQORLRQPSCKLQ--RLQL--VSCGLTSCCCODLASVLSAS 640
QY 545 PALTTLDLSGCGLPAPMVTYLCVAVLHQHCGGLQTLSSLASVELSEGLQELQAVKRAKPD 604
DB 641 PSLKELDLQNNLDVGVRLICEGLRHACKLRLGLDCTTLDSEMKQELRALEGEKRFQL 700
QY 605 VI 606

DB 701 LI 702

RESULT 3

AS9000

mater protein [imported] - mouse

C:Species: Mus musculus (house mouse)

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

A:Accession: AS9000

R:Tong, Z.B.; Nelson, L.M.

Endocrinology 140, 3720-3726, 1999

A:Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature

A:Reference number: AS9000; MUID:99360614; PMID:10433232

A:Accession: AS9000

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1111 <KUR>

A:Cross-references: GB:AF074018; NID:G5802697; PID:AA051762.1; PID:G5802698

C:Genetics:

A:Gene: Mater

Query Match 14.6%; Score 468.5; DB 2; Length 1111;
Best Local Similarity 23.8%; Pred. No. 2.3e-27;
Matches 178; Conservative 108; Mismatches 270; Indels 191; Gaps 15;

QY 1 MLAQPORLFILOGADELPALGQPEAAPCTDPFEAASGARVLGGLLSKALLPTALLV 59
DB 260 INSQPERLFILOGVDDMSVQLQHDMTLSRWKQDEQFIYILMYSLLRKALLPQSFLL 319
QY 60 TTRAAAPRLQGRLCSPCAEVRGFSKDKKKYFKFPRDERRAERAYFVKENETLPALC 119
DB 320 TRNTGLEKLSMVVSPYLIVBGLSASRRSQVLNENISNESDRIQVHSLIENHOLFPOC 379
QY 120 FVPFVWVICTVLRQQLGELGRDLSRTSKTTTSVYLLFTSVLSAPVADGPRLOGSLRN 176
DB 380 QAPSVCSLVCEALQQLKGRCT-PCOTLTGLYATLVFHQTLTKRPSQALSQEEQITL 439
QY 177 RMLCLAREGVJRRAGFAKELEQLERKSKVQTLFLSKKELPGVLETEVTVYQIDQSF 236
DB 440 VGLCMVAEGVWTKRSVFYDDDKYLSKESILALFHMNILLQVHNSQCYVFESHLSL 499
QY 237 QBFALALSYLLEDCGVPRTAAGGV-----GTLRGDQAPSHLVLTTRFLFG 283
DB 500 QDFFAALYVLE-----GLEMKQHPFCFIENQRSIMEVKRTDTRLLGMKRF-LG 549
QY 284 LLSAERYRDIERHFGCMVSRVKQBALRWVQGGCGGCGVAPVETEGAKGLEETSEPEE 343
DB 550 LMNKOILKTLVLFEPVPTVEQKLQHWVSLIAQOVNGTSPMD7----- 594
QY 344 EGEENYPLELLCYLCTOEDAFVROALCRPELALORVRFRCRDVAVLSYCVRCPPAG 403
DB 595 -----LDAYCLFESQDEEFVGGALKRQEVALLNQ--KMDLVSSVCLKHCQNL 643
QY 404 QMLR-----LIS-----CRLVAAQKKKK-----SLGKRLQASLGGG- 435
DB 644 KAIRVDRLDLLSDNTLCLCPVTVQETQCKPLMEWNGNFCVLSGLRNLKELDGLDSI 703
QY 436 -----SSO-----GTT----- 441
DB 704 LSQRAWKILCLELRNQSCRIOKLTFKSAEWSGLKHLKWLKLFNQNLKYLNLGNTPMKDD 763
QY 442 --KQLPASLLHP----- 451
DB 764 DMKLACEALKPKCSVETLRDSCELTIIGYEMISTILLSTRKCLSLAKNRVGVKSI 823
QY 452 -LFOAMTDPLCHSLSLTSLSHCKLPDVCRLDSEALRAAPALTELGLLHNLSEAGRLMS 510
DB 824 SLGNALSSMCLLQKLLDNCGLTPASCHLVLSAFLSNQNLTHLCLSNNSLGTGEGVQLC 883
QY 511 EGSAMPQCRVQTV---RVOLPDQRLGQYLVGMLRQSPALTTLDLSGCGQLPAPMVTYLC 567
DB 884 QFLRNPICALQRLILNHGNCIVDDAYG--FLANRLANNTKLTHLSJTMNPVGDAVKLLCE 941

Db 1098 RCPHVTAMWTFPIPFVQVHLL 11230

RESULT 5

A31857

ribonuclease inhibitor, hepatic - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 21-Jan-2000

C:Accession: A31857; A35830

R:Hofsteenge, J.; Kieffer, B.; Matthies, R.; Hemmings, B.A.; Stone, S.R.

Biochemistry 27, 8537-8544, 1988

A>Title: Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals

A:Reference number: A31857; MUID:89118268; PMID:3219361

A:Accession: A31857

A:Molecule type: protein

A:Residues: 1-456 <HOF>

R:Vicentini, A.M.; Kieffer, B.; Matthies, R.; Meyhack, B.; Hemmings, B.A.; Stone, S.R.

Biochemistry 29, 9827-9834, 1990

A>Title: Protein chemical and kinetic characterization of recombinant porcine ribonuclease

A:Reference number: A35830; MUID:91104783; PMID:2271559

A:Accession: A35830

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 82-456 <VIC>

A:Cross-references: GB:M58700; NID:G164638; PIDN:AAA63448.1; PID:G164639; GB:J02925

C:Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homolc

C:Keywords: liver

F:280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 5.7%; Score 216; DB 2; Length 456;

Best Local Similarity 27.4%; Pred. No. 1.3e-08;

Matches 77; Conservative 45; Mismatches 113; Indels 46; Gaps 7;

QY 354 ELLYCLVETO----EDAFVQALCRPPELALQVRCRMDVAVLSYCVRCPPAG----- 403

Db 17 ELLPLLOQYEVRLDDCGLTEHCCKIGSAL-----RANPSITELCLRTNELGDAGVHL 70

QY 404 --QALRLISRLVAAQEKXKSLGKRLQASLGGSSGGITTKQLPA-----SL 448

Db 71 VLQGLQSPCKI-----QKSLQNCSLTEAGCGVLPSTLRLSLPTLRELHLSNDNPLGDAG 124

QY 449 LHPFLQMTDPLCHLSLTLSHCKLPDAVCRDISEALRAAPALTELGCLLNRLNSEAGLRM 508

Db 125 LRLCEGLDPQCHLEKLOLEYCRLLTAASCEPLASVLRATRLAKELTVSNND:GEAGARV 184

QY 509 LSEGLAWPQCRVQTVVCLPDPOGL-----OYLVMGLRQSPALTTJDJSGCOLPAPMVT 563

Db 185 LGOGLASDAQLETLRL-----NCGLTPANCKOLGIVASQASLRELDLGSNGLDAGIA 240

QY 564 YLCVAVLQHQCGGLQTLASVLSVLSQSLQELQAVKRAKPD 604

Db 241 ELCPGLLSPASRLKTLWECEDITASGCRDLCRVLQAKETL 281

RESULT 6

A31858

ribonuclease-angiogenin inhibitor - human

N:Alternate names: ribonuclease inhibitor, placental

C:Species: Homo sapiens (man)

C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 19-May-2000

C:Accession: A31858; S02012; S23933; S48636; T47188

R:Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.

Biochemistry 27, 8545-8553, 1988

A>Title: Primary structure of human placental ribonuclease inhibitor.

A:Reference number: A31858; MUID:89118269; PMID:3219362

A:Accession: A31858

A:Molecule type: mRNA

A:Residues: 1-461 <LEE>

A:Cross-references: GB:M22414; NID:G186260; PIDN:AAA59130.1; PID:G307040

R:Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweiger, M.

EMBO J. 7, 4151-4156, 1988

A:Reference number: S02012; MUID:89210799; PMID:3243277

QY 568 VLQHQCGGLQTLASVLSVLSQSLQEL 594

Db 942 ALKEPTCYLQGLVLCQTLNCCEDL 968

RESULT 4

A48843

MHC class II transactivator - human

N:Alternate names: CITA

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 38-Oct-1999

C:Accession: A48843

R:Steinme, V.; Otten, L.A.; Zufferey, M.; Mach, B.

Cell 75, 135-146, 1993

A>Title: Complementation cloning of an MHC class II transactivator mutated in hereditary

A:Reference number: A48843; MUID:94006536; PMID:8402893

A:Accession: A48843

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1130 <STE>

A:Cross-references: GE:X74331; NID:G414112; PIDN:CAA52354.1; PID:G414113

Query Match 9.5%; Score 304.5; DB 2; Length 1130;

Best Local Similarity 25.8%; Pred. No. 7.2e-15;

Matches 174; Conservative 80; Mismatches 262; Indels 167; Gaps 29;

QY 1 MLAQPORLLFILDGADLPALGGPEAAPC-TDPEAASGARVLGGSLSKALLPTALLLVT 59

Db 487 ILKRPDRVLLILDAPFELEAQQDFLHSTCGPAEPSCSLRGLLAG-FQKLLRGCTLLIT 546

QY 60 TRAAAGRLGRLCSPOCA-EVRFSDKDKKKYKFFDERRAE---RAYRFVKNET 115

Db 547 ARPR--GRVQSLSKADALFELSGFMEQAQAYVMRYFESSGMTEHQDRAILLDRPL 604

QY 116 FALCFPVFCWIVCTVLRQGLEGRDLRSKTTTSVYJ-LFITSVLSSAPVADGPRLQ 174

Db 605 LSHSHGPTLCRAVCQSEALLEGED-AKLPSLTGLYVGLGRALDSDPP-----G 655

QY 175 DLNLCRLAREGVLRRAQFAEKELEQLGRGSKVOTFLSK---KELPGVLETEVTV-Q 230

Db 656 ALAEALAKLAWE--LGRHQ---STLEDQFPSSADVTWAKGLVCHPPRAESELAPFS 710

QY 231 FIDQSQEFL-AAALSYLLEDCGVPRTAAGVGTLRLGDAQPHSH-LVLTTRFLGLLSAE 288

Db 711 FLLOQFLGALWALSKEIKDFELPQVLA---LTPKXRPYDNWEGVREFLAGLIFQF 765

QY 289 RMEDIERHFGCYTSERV--KQAL-RWVGOGGCGFVAPETVTEGAKGHLEDETEEEEEE 345

Db 766 PARCJGALLGPSAAASVDRKQKVLARYLKRLOG-----TLRAQL----- 806

QY 346 GEENPYFLELLYCYTQEDAFVQALCRFP-ELALQVRFCRYDVAVLSYCVRC----- 399

Db 807 -----LELLHCAHEAEAGIHQHVQELPGRLFLGTRLTTPDDAHVLGKALEAAGQDF 859

QY 400 -----CPAG-QALRLISC-----RLVAQEK----- 419

Db 860 SLDLRSTGICPSGLSVGLSVCFTRFRAALSDTVALWESLRQHGETKLLQAAEERFT-EP 919

QY 420 -KKKS-----LGKRLQASLGGSSGGITTKQLPA-----SLHPLFCAMTDP----- 459

Db 920 FRAKSLKVEDGLKLVQVQRTSSSEDTAGELPAVRDJKKLEFALGVPVGPOAFPKAVRI 979

QY 460 -----LCHLSLTLSHCKLPDAVCRDISEALRAAPALTELGCLHNRLNSEAGLRVLSGL 513

Db 980 LTAFFSSLOHLLDLSLENKICDGVGSLSNFTFQLKSLLETNLNSQNNITDLGAYKLAEL 1035

QY 514 AWPQCRVQTVRVQ-----LPD-----PQGLQYLVGLMR 542

Db 1040 --PSLAASLLRLSYNNICICUGAESLARVLPDWSLRVMDVQYKNKFTAAGAAQLAASLR 1097

QY 543 QSPALTTLDLGGCQLPAPMVTYL 565

Dd 260 INFILLVFCLLSNE-----GSDI-----KMTQVLFMSMTREVBSSHLKGEVPLDKVGAQY 311
 QY 178 NJCLAREGVGRRAPFAEKELEQLERGSKVOTLFLSKKELPG-----VLEETVYQ-- 232
 Dd 312 KLAELAYKGLQORUKLVFEKTFDDVLADMTNFTHTYVDISSGIRIKLEGNKASVET 372
 QY 233 DQSFQFLAA-----LSYLLDGGVPRTAAGGVGTLRGRDAQPHSHLVJ-- 276
 Dd 372 HHKIOBEFYAAVYLMFLVSRIPLFQLXTIFED-----TQMSVVVKFMFGICNPHAYKQJKI 426
 QY 277 -----TTRELFGU-----LSAERNRDIERHFGCMWSERVQAEALRWVQ-----CC 316
 Dd 427 IFFATMIKQYEEKKKFLSXYWESLSAKSEDLIRFGMLHEVNDDETSKKFKCCLPVGJ 486
 QY 317 GCGCPGVAPEVTEGAKGLEDTEPEEEEEEENYPLELLYCLYE----- 361
 Dd 487 KGVGPKHLPVKDLVVAJAKSFTRPKHLRLRSNNKTTTTEVLETLRGLNGHTTTTTRFVIN 546
 QY 362 --TCEDAFVROALCRPEELALQVRFCRMDVAVLSYCVRCPCAGQALRLISRLVAQEK 419
 Dd 547 NIEMKSLMELLILLHD--ANEXLRF--DDVTNLS----- 577
 QY 420 KKSLSKELQASLGGSSQGTTKQLPASLHLPFOANTDP-----LCHSSLTSHC 471
 Dd 578 -----SYWESLSNAINORSNKIQLDLWTH--QQNDNDVVYLAGCLGN--SRLNKYSHT 627
 QY 472 KLPDAVCRDLSEALRAAPAL-----TELGLLHNLSEAG 505
 Dd 628 YISSDOCRVLKQAIQGLPSIQVQLYPDILSTVLYNMRPIIRFDNTSYVYVHDOOFESS 687
 QY 506 LRVLSEGLAWPQCRVOTVRVOLPDPQRGILQYVLMRLSPAL--TLD--SQCLPAP----- 560
 Dd 688 -----SKCWIGS-----RG-----GKLVGGCELWVPPGALE 714
 QY 561 --WVTVLCVQLHQQCGCQJOT-----LSIASVELSEQLQELQ-----AVKRPDLVZ 606
 Dd 715 KDVEIKLTASLSESEFLEPTLQCELASLTLLKKQVTKLQTHVLDLTKETIRCKVTLVY 774
 QY 607 THPALDGH 614
 Dd 775 TRVTTTVH 782
 RESULT 9
 T17233
 Hypothetical protein DKFZp4341216.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 15-Oct-1999
 C:Accession: T17233
 R:Koehrer, K.; Beyot, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, September 1999
 A:Reference number: Z18722
 A:Accession: T17233
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-914 <KOE>
 A:Cross-references: EMBL:AL117435
 A:Experimental source: adult testis; clone DKFZp4341216
 C:Genetics:
 A:Note: DKFZp4341216.1

Query Match 4.0%; Score 130; DB 2: Length 914;
 Best Local Similarity 21.7%; Pred. No. 0.11;
 Matches 133; Conservative 95; Mismatches 222; Indels 164; Gaps 33;

QY 34 EAASGARVVGGLLSKALL---PTALLLVTTTRAAP-----GRVQGRLL--CSPOCA 78
 Dd 14 EVPSGLQ-LEQLPSGLLTH1PTAGLPTSLGGGLPYCHQAWLDFRRLEALLQN-QAACA 72
 QY 79 EVRGSFSDKKKXYFYKFFEDRRAERAEAYRFVXENETLF-ALCFVFPVCHVCTVRLQQLE 137
 Dd 73 LLGGAIAS-----VKAVPQFMPEGEVGGQLLOQTEVLMQQVLDSEWLAWLQC-----Q 119

QY 138 LGRDLRSRTKTTTSVYLL--FITSVLSAPVADGPRLQGGDLRNLCLRLAREGV----LGR 191
 Dd 120 GGRLETLWIKQEVPEVTLSPDYRTAMDKADELYD--RVDTGLLHQLTQSNQRIQALELVOT 177
 QY 152 AQFAEKELQELBLRGSKVOTLFLSKKELPG---VLEETVYQFIDQSPQEFLLAALSYLE 248
 Dd 178 LEARESGLHQIEVMVLLQOYGVWPALEAGEPFLDMLLQAQGSFQELYVQAQEQVRQGEKFLQ 237
 QY 249 DGGVPRTA-----AGGVCGLLRGDAQPHSHLVLTTRFLFLGILLSAERMDIERHFCM 300
 Dd 238 -----PLTGWAAAEIDPPGARFLAURAQLTEFSR-ALAQ-----CORTADAERLF-- 282
 QY 301 VSERVKQFALRWVQGGCGCPGVAPEVTEGAKGLEDTEPEEEEEEENYPLELLYCLY 360
 Dd 283 --QLFREALTMAE-----EGORVLALEEQ-----BRPGVVLLQGLQ-LH 317
 QY 361 ETQEDAFVROALCRPEELALQVRFCRMDVAVLSYCVRCPCAGQALRLISRLVAQEK 419
 Dd 318 WT-----RHPD--LPPAHFRKM-----WALATGLGSEAIRO-ECRWAWARQD 357
 QY 420 KKSLSKELQASLG---GSSSQGTTKCLPASLHP-----LFOAMTDPCLHL- 463
 Dd 358 TWALDQKLEASLKLPPVGVSTASLCVQVPAAPPAHPPLRKAYSTDRNLGQSLSEFACHCH 417
 QY 464 SSSLTSHCKLPDAVCRDLSEALRAAP-----ALTELGLLHNLSEAGLRLMSEGLAW 515
 Dd 418 HAATIAACRRPEAGGALPQASPTVPPGSSDPRSLNRLQL-----VLAEMVAT 466
 QY 516 PQCRVOTVRVOL-----PDQRGJQ--YLVGMRLQ-----SPALTTLD--SGGQLP 558
 Dd 467 EREYVRALEYTMENYFPPELDRPDPVQGLRGORAHLFGNLEKLRDFHCHFFLRELEACTRH 526
 QY 559 APWTVYLCVQLHQ 572
 Dd 527 PPRVAY--AFURHR 538
 RESULT 10
 S75472
 GTP-binding protein hflX - *Synechocystis* sp. (strain PCC 6803)
 N:Alternate names: protein sir1521
 C:Species: *Synechocystis* sp.
 A:Variety: PCC 6803
 C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #ext_change 02-Feb-2001
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.
 C. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu
 DNA Res. 3, 109-136, 1996
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*.
 A:Reference number: S74322; XUID:97061201; PMID:8905231
 A:Accession: S75472
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-534 <KAN>
 A:Cross-references: EMBL:D90911; GB:AB001339; NID:G1633083; PIDN:BAAI8033.1; PID:d1018
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Genetics:
 A:Gene: hflX
 A:Start codon: GTG
 C:Superfamily: translation elongation factor Tu homology
 C:Keywords: GTP binding; nucleotide binding; P-loop
 F:362-489/Domain: translation elongation factor Tu homology <ETU>
 F:368-375/Region: nucleotide-binding motif A (P-loop)
 F:486-489/Region: GTP-binding NKXD motif
 Query Match 3.6%; Score 116; DB 2: Length 534;
 Best Local Similarity 23.6%; Pred. No. 0.63;
 Matches 110; Conservative 60; Mismatches 170; Indels 126; Gaps 25;

QY 182 LAREGVL-----GRRAPFAEKEL--EQELRGSKVOTLFLSKKELPGVLETEVYQFIDQS 235
 Dd 140 LAEEAILDVLSELEAEF-EREFTARQVEAGQERVLMGLATSELD-----DQS 186

QY 236 FOEFFAAJLYLLEDDGV-----PRTRAA--GGVOTLLRGDAQPHSHLVLTTR 279
DB 187 FAEGITELSERJVDSDAQAQVJVMQNRKSRPHQTVVGEKVEELALAVQTTGANSILVFDK 246
QY 280 FLFGLLSAERMDIERHFCQWSEK-----OEALRWVQG 315
DB 247 D-----LSAAQVRNLEQRCURVIDRTELILDIIPAQASRAGKQVLELAQLEYLLPKLVG 302
QY 316 QGCGCGVGAPEVTEGAGJEDTEPEEEBEGEPNYPLELLYCLYETQEDAFVQALCRF 375
DB 303 RGCGMSRLGGGI--GTRPGCTKLETRNTIQSIKALQ-----KQVNELOSHR 349
QY 376 PELALQVRFCRMDVAVLSYVRCPCAGQAL---RLISCRVAAQEKKKSLGKRLQASL 432
DB 350 SRLRNORQQAQVPTVAIVG---TNAGKSTLLNALTQADIYAADO-----LFATL 396
QY 433 GCGSSGTTKQLPASLHLPLFQAVTDLPLCHLSLTLSHCKLPDAVCRDLSEALRAA-PAL 491
DB 397 -----DPTTRJC--SLLDPENQTY-HPIJLTDTVGFH-KLPDA---LVDAFRATLBEV 443
QY 492 TELGLLHNLRSAGRLMS-EGLAWFO--CRYQTVRVQLPDPQRGLQYLVGMRLQSPAL 548
DB 444 TEADLL-----LQWDLSDRAMRQRIASVANILAEMPLATAPMVMVFNKIDQVPSA 495
QY 549 TLDLSCCQLPAPMVTYLCVAVLQHGCGGLQTLASVELSEQSLQEL 594
DB 496 ----LAAQAQYPEAIFLAAA---QGIGLETL---KKRLLEQLITDL 532

RESULT 11
T35861
probable large secreted protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35861
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, April 1999
A:Reference number: 221591
A:Accession: T35861
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-877 <SAL>
A:Cross-references: EMBL:ALC49727; PIDN:CAB41562.1; GSPDB:GN00070; SCCEB:SC9B.116C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCCEB:SC9B.116C

Query Match 3.6%; Score 114.5; DB 2; Length 877;
Best Local Similarity 22.0%; Pred. No. 1.5;
Matches 113; Conservative 63; Mismatches 174; Indels 163; Gaps 27;
QY 114 TLPALCFVFPV---CMIVCTV-----LRQQLGELGDLSTKTTTSTVLLFITSVL 162
DB 21 SLLALVGVVLAQVGVTTIQMSQGLRLRAQQLSDTG--AVGTZALAQKERSL 78
QY 162 SSAPVA--DQPLQGDRLNCLAREGLVGRRAQFAEKEL-----EQELRGSKVQTLFSL 215
DB 79 SAALWLAAPDGRSELDAQR-----KXTDAAVAGLVQSDAIESAPTRVSDRLYS 127
QY 216 KLELPVLE-----TEVT-YQFIDQ-----SFOFLAAV-SYLLDDGVPT 255
DB 128 VLGSVGSLEYRNQVDDPSDITAEALQDQYTSIVDEQIHAFQF-----LSQ-VDDGDLTSQ 182
QY 256 AAGGVCTLLRGDAQPHSHLVLTTRFLGLLSAERMDIERHFCQWSEKVEALRW-VQ 314
DB 183 AGPLVA:EHNAELVSCEDALTLAWPSGRMDEQW-----SRFAQLVHTR-----RWLVQ 232
QY 315 GQCGCGVGAPEVTEGA-----KGLEJTE-----EPBEEBEGEENFPY 352
DB 233 DQ-----VPSLTGSAKQJTERILASSEWKSQVQEDQVLEARSAGRGDRIDLPDARQ 286
QY 353 ----LELLYCLYETQEDAFVRQALCRFPPELALQVRFCRMDVAVLSYVRCPCAGQALRL 408

DB 287 WITAFEKI---STQYQGLIRQOTDGLDRSABEARGLLIKAGVLS-----AGGLIAL 335
QY 409 ISCRVAAQEKKKSLGKRLQASLQSGSSQGTTKQLPASLHLPLFOAMTDP-CHLSSITL 468
DB 336 LJC--IVMSWRITRSRLRG-----LRLATLSL 363
QY 469 SHCKLPDAVCR---DLSEALDAAPAL-----TELGLLHNLRSAGRLMSLSEGLAMPQCRV 520
DB 364 AEERLPDVVARLERGETVDVESATPLOYGRDELQOVAQAFNA-----CRTAV 412
QY 521 QTVRVQLPDPQGLQ-YLVGMRLQSPALTTLDL 552
DB 413 HTA-VELADTRRGFKVILGIARQSONLVNMOL 444

RESULT 12
T31668
hypothetical protein COSL1.5 - sea squirt (Ciona intestinalis)
C:Species: Ciona intestinalis
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Jun-2000
C:Accession: T31668
R:Bird, A.P.; Clark, V.; Jones, S.J.; Leighteb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z21050
A:Accession: T31668
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1075 <BIR>
A:Cross-references: EMBL:Z80904; PIDN:CAB02589.1
C:Genetics:
A:Introns: 48/1; 557/3; 611/3; 685/2; 721/2; 739/1; 779/3; 820/3; 853/2; 914/2; 993/1

Query Match 3.5%; Score 114; DB 2; Length 1075;
Best Local Similarity 19.0%; Pred. No. 2.2;
Matches 126; Conservative 102; Mismatches 278; Indels 158; Gaps 27;
QY 41 VLGGLLSKALLPTALLVTTRAAAPRQGRCSQCAEVRGFSQCKKQKYPKFRFR 100
DB 198 IMKNLNLTLPAKXIVTSTLHQMYKLPDVPRTSIFEVLGLLEEAKNLTQCGEKY 257
QY 101 RAERAYRVKENETLPAFCVFP--FVCWIVCTVLSQQLLEGRDLSRTSKTTTSVYLLFIT 158
DB 258 PAIK--KILQCPNLHAHLCYLPINFILIVFCLLSNE---GSDI-----KTMTOVLIFSMT 307
QY 159 SVL-----SSAPVADGPRLOGLDRLNCLAREGLVGRRAQFAEKELQELRGSKVQTL 212
DB 308 RFVEUSHLKGEVPL---DKVGAEMVKLARAYKGCQKLVFEKTFDQVLADEMTNF 364
QY 213 FLSKKE-LPG-----VLETEVTYQFIDQSFQEFLLA-----LSY-----LLEDGGVP 253
DB 365 PHTYVDKSSGIEMKILEGNKRSYFTHLIQWEPYAAVY-MLFVSYPREFBOLKPIFKDAQWK 424
QY 254 RTAAGVGVTLLRGDAQPHSHLVLTTRF-----LFGLLSAERVRDIER 295
DB 425 RV-----VGMPF-GICNPAYKQLKLVFFATMIKDYEEKKELMVPMMESLASARGEDLIR 479
QY 296 HFCQWSEKVEALRWVQSGCGPGVAPEVTEGAKGLED-----TEEPBEEBEGEE 349
DB 480 RFG-----WLHEYND-----DESKKFEDYLPVGLKWDAPKHLSEVKD 517
QY 349 PNY-----PLELYCLYETQEDAFVRQALCRFPPELALQVRF-----CRMVAVLSYCV 397
DB 518 LVVALKSFTKPKHRLDSYETTTEVLETLRGVHGHTTTTITRFINNIEMKDSLMEILL 577
QY 398 RCCPAGQALRLISCRVAAQEKKKSLGKRLQASLQSGSSQGTTKQLPASLHLPL----- 452
DB 578 LHLDAEMEELFRVNTLSPYMER-----LSNAINQFNSNKIQLVLKGLKHLDDYD 626
QY 453 FOAMTDPCHLSLTLSHCKLPDAVCRDLSEALRAAPAL-----TELGLLHNLRL 501
DB 627 VKYLACCLGNISLLYMWGTDISSDQCSVLKQAIQOLPSIQVHQLYDPDILSTYLVNARNWI 686

502 SEAGRLWSEGLAWPQCRVQVTRVQLPDPQRGLQYLVLGMLRQSPALITTLDLSCQLPAP- 560
587 SN-----WFVSEVQVPEAL-NLGNGLKDGSGQPRRNEELCSQDVS---VPAPT 731
561 NVTY--LCAVLQHQCGLCTLISLA-----SVELSEQSLCELOAVKRAKPDVLVITHPAL 611
732 QVNYNHYNCAPIYHQ---VRFLTAQFILFMINSUNLHPNGLDQVEKWRUEDESLEYLTPKJ 788
612 DCHP 615
789 QCEP 792

RESULT : 3
AF0697
probable type III secretion protein ssav [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2302
C:Accession: AF0697
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Parrar,
S.; Moulie, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:2534947; PMID:11677608
A:Accession: AF0697
A:Status: Preliminary
A:No-scale type: DNA
A:Residues: 1-681 <PAR>
A:Cross-references: GB:A133382; PIDN:CAD01951.1; PID:g16502793; GSPDB:GNC0176
C:Genetics:
A:Gene: ssav
C:Superfamily: regulatory protein lcrD

Query Match 3.5%; Score 113.5; DB 2; Length 681;
Best Local Similarity 21.8%; Pred. No. 1.3;
Matches 143; Conservative 86; Mismatches 243; Indels 183; Gaps 33;

QY 36 ASGARVLGGJLSKALLPTALL-----LVTTTRA-----AAPGR----LQGRLC 73
DB 110 AFGKFVVGKLTVLGVVFTTITVQFVITKGIHVAEVSARFSLDGMPOKQMSIDGDLR 169
QY 74 SPQCAEVRGFSDDKKYFYKFRDERP---AERAVRFYKENETL--FALCFVPVVCWI 127
DB 170 A-----GVYDADHARTLRHQVQESRFILGADGAKFKVK-GDTAG:IVV-VNI:GGI 221
QY 128 VCTVLRQOELGRDLRSKTSKTTTSVYLLFTSVLSSAPVADGPRLOGLDLNLCRLAREGV 187
DB 222 IIAIVQY-----DMSSEAVHT-----YSVLS-----IGDG--LCQIFSLJISLSAGI 263
QY 188 L-----GRRAPAEKELEQLSGSKVCTILFLSKKEJ-----PGVLETVTYQFIDQSF 236
DB 264 IVTRVPGEKQNIATELSSQ:AR--QPQSUILFAVLMMLLALIFG-----EPFITLAF 314
QY 237 QEFLLAALSYLELDGGVPRTAGGVGTLLRGDAPSHSLVLTTRFLFGLLSAERMDIERH 296
DB 315 FSALLALPILIRRKKSVMVANGIAPEKDSMVPGA-CPLILRLTFLHSADLIRI---- 370
QY 297 FGCWVSERVQKEALRWVQGGQCGPGVAPETVTEGAKGJEDTEPEBEE---GEENPYPL 353
DB 371 -----DAMRWFLPDTGVP--LPEV-----NIEVLPEPTEKLTLLVLYOEPVFSL 412
QY 354 ELLYCLVETQEDAFVROA-----LCRFPFELALQVRFRMDY----- 390
DB 413 SI-----PAQADYLLIGADASVVGDSQTLPNGMGQICWLKMDMAHKAQGFGLDVFAQSQR 467
QY 391 -AVLSYCVRCPCPAGQARLLISCR-LVAAQEKKKKSLGKR-CAS:GGSSSGTITKQLPASL 448
DB 468 ISALLKCVLRHMGFEITGVQETRYLMMANRKNYSVELVKELQ-----RQIPINK 515
QY 449 LHPLFQAMTDFLCHLSLTLSHCKLPLDAVCRD-----LSEALRAAPALTELGLLHNRLE 503

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2003, 03:41:16 ; Search time 696 Seconds
(without alignments)
10429.294 Million cell updates/sec

Title: US-10-029-347-1

Perfect score: 2689

Sequence: 1 cggacgctggcgcgagc.....aaaaaaaaaaaaaaaaaa 3689

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 13497190.7 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.:

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2689	100.0	2689	24	DNA encoding human
2	2626	97.7	3365	22	Human G-protein cc
3	2605	96.9	2763	24	Human leucine-rich
4	2020.2	75.1	2054	24	Human leucine-rich
5	1938.6	72.1	2109	24	Human MOV18a cDNA.
6	151.2	58.4	5556	24	Pyrim domain conta
7	1571.2	58.4	6461	22	Human ATLAS-2-enco
8	1368.6	50.9	1795	22	Human vasopressin

9	832.8	31.0	2251	14	AAQ38723	ATI/AVP2 receptor
10	798.6	29.7	933	22	AAK83113	Human immune/haema
11	740.6	27.5	989	22	AAK62741	Human immune/haema
12	729.6	27.1	990	24	ABL90582	Human polynucleoti
13	464.6	17.3	704	24	ABQ29970	Oligonucleotide fo
14	464.6	17.3	704	24	ABQ29971	Oligonucleotide fo
15	432.6	16.1	704	24	ABQ29972	Oligonucleotide fo
16	432.6	16.1	704	24	ABQ29973	Oligonucleotide fo
17	427.6	15.9	678	25	ACA56413	Norway rat signall
18	289.2	10.8	463	22	AAK88542	Human digestive sy
19	277	10.3	3186	24	AAAL4363	Human PYRIN-8 cDNA
20	273.4	10.2	3300	24	AAAL47129	Pyrim domain conta
21	258	9.6	314	24	ABS63488	Human leucine-rich
22	236.6	8.8	257	25	ACA55610	Norway rat signall
23	226.4	8.4	3857	22	AAAL4323	Human PYRIN 1 (PVR
24	226.4	8.4	3857	25	ABX93556	Huma cDNA encoding
25	195.4	7.3	2847	22	AAAS01487	Human secreted pro
26	195.4	7.3	2847	25	ABZ73494	Secreted protein-e
27	193.6	7.2	1557	24	AAAL47143	Pyrim domain conta
28	179.2	6.7	3431	22	AAAL4322	Human nucleotide b
29	179.2	6.7	3431	25	ABX93555	Huma cDNA encoding
30	179	6.7	4931	24	ABL59333	Nucleotide sequenc
31	177.6	6.6	3189	22	AAI67185	Nucleotide sequenc
32	177.6	6.6	3531	24	AAAL47128	Pyrim domain conta
33	177.6	6.6	4787	23	AAAS92566	DNA encoding novel
34	171.6	6.4	3085	23	AAAS68111	DNA encoding novel
35	171.6	6.4	3150	22	AAI67184	Nucleotide sequenc
36	171.6	6.4	3263	24	AAI70684	Human nucleotide b
37	171.6	6.4	3459	24	ABQ78049	Human CGDD encodin
38	170	6.3	3069	24	AAAL47134	Pyrim domain conta
39	167	6.2	3368	24	AAAL44366	Human PYRIN-3 cDNA
40	165.4	6.2	2575	24	AAI70683	Human nucleotide b
41	165.4	6.2	2767	25	ABT16018	NOVX related polyn
42	165.4	6.2	3172	24	ABN99366	Human secreted pro
43	163.8	6.1	969	22	AAK78135	Human immune/haema
44	163.8	6.1	969	22	AAK78136	Human immune/haema
45	162.2	6.0	969	22	AAK78134	Human immune/haema

ALIGNMENTS

RESULT 1
ABS63485
ID ABS63485 standard; cDNA; 2689 BP.
XX
AC ABS63485;
XX
XX 15-NOV-2002 (first entry)
XX
XX DNA encoding human leucine-rich repeat small intestine I (HLRRS1).
XX
XX Human; human leucine-rich repeat small intestine I; HLRRS1; asthma;
XX proliferative disorder; gastrointestinal disorder; renal disorder;
XX neural disorder; reproductive disorder; calcium regulation; apoptosis;
XX immune system; anaemia; human immune deficiency virus; HIV; cancer;
XX blood coagulation disorder; autoimmune disorder; allergic reaction;
XX inflammatory condition; cardiovascular disorder; ischaemia;
XX neurological disorder; infectious disease; cytokine production;
XX expressed sequence tag; EST; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200261086-A2.
XX
XX 08-AUG-2002.
XX
XX 20-DEC-2001; 2001WO-US49739.
XX
XX 22-DEC-2000; 2000US-257774P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX

Feder J, Ramanathan C, Mintier G;

WPI: 2002-619252/66.

P-PSDB: ASG78454.

XX New isolated nucleic acid molecules encoding HLRRS11 polypeptides, or
PT their fragments and homologues, useful for preventing, treating and
PT ameliorating medical conditions, e.g. proliferative, gastrointestinal,
PT or renal disorders

XX Claim 1; Figure 1; 336pp; English.

XX The invention relates to isolated nucleic acid molecules (if) encoding
CC human leucine-rich repeat small intestine I (HLRRS11) polypeptides.
CC The nucleic acid molecules and polypeptides are useful for preventing,
CC treating and ameliorating medical conditions, such as proliferative,
CC gastrointestinal, renal, neural, or reproductive disorders, or disorders
CC related to aberrant calcium regulation or apoptosis modulation, either
CC directly or indirectly. They are also useful for treating, preventing
CC and/or diagnosing diseases, disorders and/or conditions of: immune system
CC by activating or inhibiting the proliferation, differentiation, or
CC mobilisation of immune cells; haematopoietic cells e.g. thrombocytopoenia,
CC anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
CC virus (HIV) infection, HTLV-ELV infection; blood coagulation disorders,
CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
CC myasthenia gravis; asthma or allergic reactions; inflammatory
CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
CC ischaemia; aneurysms; neurological disorders, e.g. Alzheimer's disease,
CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
CC pneumonia, or viral, bacterial, and fungal infections. The HLRRS11
CC polypeptides are useful for modulating cytokine production, antigen
CC presentation, or other processes such as boosting immune responses.
CC AB563495-AB563504 represent HLRRS11 coding sequences and PCR primers of
XX the invention.

XX Sequence 2689 BP; 534 A; 861 C; 644 G; 450 T; 0 other;

Query Match 100.0%; Score 2689; DB 24; Length 2689;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CGGACGGCTGGCGCGGAGCTGGCTGACCTGATCTCTGGACAGTCCCGGACCGCGCG 60
Db 1 CGGACGGCTGGCGCGGAGCTGGCTGACCTGATCTCTGGACAGTCCCGGACCGCGCG 60
Oy 61 CGCGGTGCGCGAGATGCTTGGCGCGGAGCTGGCTGATCTCTGGACAGTCCCGGACCG 120
Db 61 CGCGGTGCGCGAGATGCTTGGCGCGGAGCTGGCTGATCTCTGGACAGTCCCGGACCG 120
Oy 121 ACAGAGTCCCGCGGCTGGCGGCGCGGAGCGCGGCTGGACAGACCGCTTGGAGCGG 180
Db 121 ACAGAGTCCCGCGGCTGGCGGCGCGGAGCGCGGCTGGACAGACCGCTTGGAGCGG 180
Oy 181 CGAGCGGCGCGGCTGGCTAGCGGGGCTGCTGAGTAAAGGCGCTGCTGGACCGCGGCTCC 240
Db 181 CGAGCGGCGCGGCTGGCTAGCGGGGCTGCTGAGTAAAGGCGCTGCTGGACCGCGGCTCC 240
Oy 241 TGCTGGTGAACACCGCGGCTGGCGGCGCGGAGGCTGGAGGCGCGGCTGGCTGGCGG 300
Db 241 TGCTGGTGAACACCGCGGCTGGCGGCGCGGAGGCTGGAGGCGCGGCTGGCTGGCGG 300
Oy 301 AGTGGCGGAGGTGGCGGCTTCCGACAAAGACAAAGAGAGTATTTCTACAAGTTCT 360
Db 301 AGTGGCGGAGGTGGCGGCTTCCGACAAAGACAAAGAGAGTATTTCTACAAGTTCT 360
Oy 361 TCCGGGATGAGAGAGGCGCGGCTTACCGCTTGGTGAAGGAGAGAGAGAGAGTCT 420
Db 361 TCCGGGATGAGAGAGGCGCGGCTTACCGCTTGGTGAAGGAGAGAGAGAGAGTCT 420
Oy 421 TCGCGGTGTGCTTGTGCGCTTGTGCTGGATTCGCTGTCACCGTGGTGGCGGCGGAGC 480
Db 421 TCGCGGTGTGCTTGTGCGCTTGTGCTGGATTCGCTGTCACCGTGGTGGCGGCGGAGC 480

Oy 481 TGGAGCTCGGTGGGACCTGTGCGCGACGTCCAAAGACCCACACCGTCACTGTACCTGCTTT 540
Db 481 TGGAGCTCGGTGGGACCTGTGCGCGACGTCCAAAGACCCACACCGTCACTGTACCTGCTTT 540
Oy 541 TCATCACCAGCGTTCGTAGCTCGGTCCGGTAGCGAGCGGCGCGCGGTTCAGGCGGAC 600
Db 541 TCATCACCAGCGTTCGTAGCTCGGTCCGGTAGCGAGCGGCGCGCGGTTCAGGCGGAC 600
Oy 601 TGGCGAATCTGTGCGGCTGTGCGCGGAGGCGTCTCGAGCGGAGGCGGAGTTCGCG 660
Db 601 TGGCGAATCTGTGCGGCTGTGCGCGGAGGCGTCTCGAGCGGAGGCGGAGTTCGCG 660
Oy 661 AGAAGGAACTGGAGCAACTGGAGCTTCGTGCTCCAAAGTGCAGACGCTGTTCTCAGCA 720
Db 661 AGAAGGAACTGGAGCAACTGGAGCTTCGTGCTCCAAAGTGCAGACGCTGTTCTCAGCA 720
Oy 721 AAAAGGAGCTGGCGGCTGTGGAGACAGAGGTCACTACCACTTCATCGACAGAGCT 780
Db 721 AAAAGGAGCTGGCGGCTGTGGAGACAGAGGTCACTACCACTTCATCGACAGAGCT 780
Oy 781 TCCAGGAGTTCCTCGCGGCACTGTCTACCTGCTGGAGACGCGGCGTGCACGAGCG 840
Db 781 TCCAGGAGTTCCTCGCGGCACTGTCTACCTGCTGGAGACGCGGCGTGCACGAGCG 840
Oy 841 CGGCTGGCGGCTTGGAGACCTCTGCTGGAGACGCGCGCACAGCCACAGCCACTTGTGTC 900
Db 841 CGGCTGGCGGCTTGGAGACCTCTGCTGGAGACGCGCGCACAGCCACAGCCACTTGTGTC 900
Oy 901 TCACCAAGCGCTTCCTTCCTGGAGCTGTGAGCGGAGCGATGCGGACATCGAGCGCC 960
Db 901 TCACCAAGCGCTTCCTTCCTGGAGCTGTGAGCGGAGCGATGCGGACATCGAGCGCC 960
Oy 961 ACTTCGGCTGCAATGCTTCAGAGCGTGTGAAGAGGAGGCGCTCGGTGGTGGAGGAC 1020
Db 961 ACTTCGGCTGCAATGCTTCAGAGCGTGTGAAGAGGAGGCGCTCGGTGGTGGAGGAC 1020
Oy 1021 AGGACAGGCTGCCCGGAGTGGCACAGAGGTGACCGAGGCGCGGCGGCGGCGGCGG 1080
Db 1021 AGGACAGGCTGCCCGGAGTGGCACAGAGGTGACCGAGGCGCGGCGGCGGCGGCGG 1080
Oy 1081 ACACCAAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Db 1081 ACACCAAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Oy 1141 TGTACTGCTGTACAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 1141 TGTACTGCTGTACAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Oy 1201 CGGAGCTGGCGCTGCGAGGAGTGGCGCTTCTGCGCATGAGCGTGGCTGTTCTGAGCTACT 1260
Db 1201 CGGAGCTGGCGCTGCGAGGAGTGGCGCTTCTGCGCATGAGCGTGGCTGTTCTGAGCTACT 1260
Oy 1261 GGTGAGGTGCTGCTGCTGCAAGGAGCTGCGGCTGATCAGCTGAGATGCTGCTGCTG 1320
Db 1261 GGTGAGGTGCTGCTGCTGCAAGGAGCTGCGGCTGATCAGCTGAGATGCTGCTGCTG 1320
Oy 1321 CGCAGGAGAGAGAGAGAGAGAGCTTGGGAGAGCGGCTCCAGGCGAGGCTGGGTGGCGCA 1380
Db 1321 CGCAGGAGAGAGAGAGAGAGAGCTTGGGAGAGCGGCTCCAGGCGAGGCTGGGTGGCGCA 1380
Oy 1381 GTTCTCAAGGACCAAAACAACTGCCAGGCTCCCTTCTTCATCCACTCTTTTCAGGCAA 1440
Db 1381 GTTCTCAAGGACCAAAACAACTGCCAGGCTCCCTTCTTCATCCACTCTTTTCAGGCAA 1440
Oy 1441 TGACTGACCACTGTGCGCTCTGAGAGCGCTCAGCTGTCCTGCTGCAAACTCCCTGAGC 1500
Db 1441 TGACTGACCACTGTGCGCTCTGAGAGCGCTCAGCTGTCCTGCTGCAAACTCCCTGAGC 1500
Oy 1501 CGGCTTGGCGAGACCTTTCTGAGGCGCTTGGAGGAGCGGCGGCGGAGCTGACGAGCTGGG 1560
Db 1501 CGGCTTGGCGAGACCTTTCTGAGGCGCTTGGAGGAGCGGCGGCGGAGCTGACGAGCTGGG 1560

1561 TCTCCACAAAGGCTCAGTGGAGCGGAGCTGCGTATGCTGAGTGAGGCGCTAGCGCTGGC 1620
1562 TCTCCACAAAGGCTCAGTGGAGCGGAGCTGCGTATGCTGAGTGAGGCGCTAGCGCTGGC 1620
1621 CGCAGTGCAGGCTGCAGAGCTCAGGCTACAGCTGCTGCTGAGCGGCTGCGCTGCGCT 1650
1622 CGCAGTGCAGGCTGCAGAGCTCAGGCTACAGCTGCTGCTGAGCGGCTGCGCTGCGCT 1650
1691 ACCTGGTGGTATGCTTGGGAGAGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1740
1692 ACCTGGTGGTATGCTTGGGAGAGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1740
1741 AACTGCCCGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1742 AACTGCCCGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1801 TGCAGACCTCAGTCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1802 TGCAGACCTCAGTCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1861 CTGTGAAGAGCAAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
1862 CTGTGAAGAGCAAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
1921 AACTCCCAAGAACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
1922 AACTCCCAAGAACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
1981 ACCTTAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
1982 ACCTTAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
2041 TGGGCTTTGGGAACTTTTGGGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2042 TGGGCTTTGGGAACTTTTGGGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2101 ACGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
2102 ACGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
2161 ACCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
2162 ACCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
2221 CGATTCCCTTGAAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
2222 CGATTCCCTTGAAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
2281 TTAACCTAAGCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
2282 TTAACCTAAGCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
2341 GTGGCCCAACAGCCCTCAGAACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2342 GTGGCCCAACAGCCCTCAGAACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2401 CTCCCTGTTAGGAATGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
2402 CTCCCTGTTAGGAATGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
2461 TTTGAGATCTTTGGGAGGAGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2462 TTTGAGATCTTTGGGAGGAGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2521 GTTGGGAGCATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
2522 GTTGGGAGCATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
2581 TGTGGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
2582 TGTGGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
2641 GTGCTCACAAAAAAGGAG 2689

Db 2641 GTGCTCACAAAAAAGGAG 2689
RESULT 2
AAD08836
ID AAD08836 standard; cDNA; 3365 BP.
XX AAD08836;
XX C4-SEP-2001 (first entry)
XX Human G-protein coupled receptor-2 (GCRC-2) cDNA.
XX Human; G-protein coupled receptor-2; GCRC-2; gene therapy; cirrhosis;
XX transgenic animal; proliferative disorder; actinic keratosis; hepatitis
XX nephropathy; cancer; breast; bladder; bone marrow; brain; uterus;
XX leukemia; adenocarcinoma; lymphoma; melanoma; myeloma; epilepsy; stroke;
XX neurological disorder; Alzheimer's disease; Parkinson's disease; nausea;
XX Huntington's disease; multiple sclerosis; dementia; angina pectoris;
XX central nervous system disorder; cardiovascular disorder; hypertension;
XX atherosclerosis; congestive heart failure; gastrointestinal disorder;
XX dysphagia; peptic esophagitis; spasm; gastritis; anorexia; pyrosis;
XX pancreatitis; Crohn's disease; diarrhoea; autoimmune disorder; anaemia;
XX inflammatory disorder; Acquired Immune Deficiency Syndrome; AIDS;
XX Addison's disease; allergy; asthma; diabetes mellitus; antithyroid;
XX atopic dermatitis; glomerulonephritis; Grave's disease; psoriasis;
XX rheumatoid arthritis; ulcerative colitis; osteoporosis; antifungal;
XX metabolic disorder; obesity; nontropic; protozoicide; virucide; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..2673
XX /*tag= a
XX /product= "Human GCRC-2 protein"
XX WC20C142288-A2.
XX 14-JUN-2001.
XX 07-DEC-2000; 2000WO-US33382.
XX 10-DEC-1999; 99US-0172852.
XX 22-DEC-1999; 99US-0171732.
XX 14-JAN-2000; 2000US-0176148.
XX 21-JAN-2000; 2000US-0177331.
XX {INCY-} INCYTE GENOMICS INC.
XX Burford N, Baughn MR, Au-Young J, Yang J, Lu DAM, Reddy R;
XX WPI: 2001-381635/40.
XX P-PSDB; AAE04546.
XX New human G-protein coupled receptor polypeptides for diagnosing,
XX preventing, and treating cell proliferative, neurological,
XX cardiovascular, gastrointestinal, autoimmune and metabolic disorders
XX Claim 5; Page 159-160; 175pp; English.
XX The present sequence is human G-protein coupled receptor-2 (GCRC-2)
XX cDNA. GCRC-2 is useful in somatic or germline gene therapy to correct a
XX genetic deficiency, to express a conditionally lethal gene product and
XX to express a protein which affords protection against intracellular
XX parasites and also for diagnosis of disorders associated with expression
XX of GCRC-2. GCRC-2 is also useful for generating hybridisation probes useful
XX in mapping the naturally occurring genomic sequences and to create
XX knockin humanised animals (pigs) or transgenic animals (mice or rats) to
XX model human diseases. GCRC-2 is used to diagnose, prevent and treat
XX proliferative disorders (actinic keratosis, arteriosclerosis, cirrhosis,
XX hepatitis and cancer); cancer (breast, bladder, bone marrow, brain,
XX uterus cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma)

CC	neurological disorders (epilepsy, stroke, Alzheimer's, Huntington's, Parkinson's disease, multiple sclerosis, dementia and other central nervous system disorders); cardiovascular disorders (angina pectoris, hypertension, atherosclerosis, congestive heart failure); gastrointestinal disorders (dysphagia, peptic esophagitis, esophageal spasm, gastritis, gastric carcinoma, anorexia, nausea, abdominal angina, pyrosis, pancreatitis, Crohn's disease, diarrhea); autoimmune/inflammatory disorders (acquired immunodeficiency syndrome (AIDS), Addison's disease, allergies, anaemia, asthma, diabetes mellitus, acropia, dermatitis, glomerulonephritis, Grave's disease, osteoarthritis, psoriasis, rheumatoid arthritis, ulcerative colitis, bacterial, fungal, parasitic, protozoal and helminthic infections) and metabolic disorders (obesity, osteoporosis, viral infections).
XX	Sequence 3365 BP: 612 A; 1106 C; 1114 G; 533 T; 0 other;
SC	Query Match 97.7%; Score 2626; DB 22; Length 3365; Best Local Similarity 99.8%; Pred. No. 0; Matches 2629; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY	7 CQTGGGCGCGAGCTGGCTTACCTGATCTCTGACCACTGCCCGACCGCGCGCGCG 66
DB	731 CGGGCAGCGCAGCGCTGGCTTACCTGATCTCTGACCACTGCCCGACCGCGCGCGCG 790
QY	67 TCCCGCAGATCTTGGCCAGCGCAGCGGCTGCTTTCATCTCTGACCGCGCGACGAC 126
DB	791 TCCCGCAGATCTTGGCCAGCGCAGCGGCTGCTTTCATCTCTGACCGCGCGACGAC 950
QY	127 TCCCGCGCTTGGGGGCGCTGAGCGCGCGCTTGCACAGACCGCTTCTGAGCGCGACG 186
DB	851 TCCCGCGCTTGGGGGCGCTGAGCGCGCGCTTGCACAGACCGCTTCTGAGCGCGACG 910
QY	187 GCGCGCGGCTGCTAGCGCGGCTGCTGAGTAAAGCGCTGCTGCCCGCGCGCTTCTGCG 246
DB	911 GCGCGCGGCTGCTAGCGCGGCTGCTGAGCAGCGCGCTGCTGCCCGCGCGCTTCTGCG 970
QY	247 TGACACAGCGCGCGCGCGCTTCCGGAGAGGCTGAGCGCGCGCTGTTTCCCGCAGTGG 306
DB	971 TGACACAGCGCGCGCGCGCTTCCGGAGAGGCTGAGCGCGCGCTGTTTCCCGCAGTGG 1030
QY	307 CCGAGTCCGGCGCTTCTCTGACAAAGCAGAAAGAAAGTATTTCACAGTCTTCCGGG 366
DB	1031 CCGAGTCCGGCGCTTCTCTGACAAAGCAGAAAGAAAGTATTTCACAGTCTTCCGGG 1090
QY	367 ATGAGAGAGGCGCGAGCGTCCCTACCGCTTCTGAGAGAGAGCAGACGCTGTTCCGGC 426
DB	1091 ATGAGAGAGGCGCGAGCGCGCTTACCGCTTCTGAGAGAGAGCAGACGCTGTTCCGGC 1150
QY	427 TGTGTTCTGTCGCTTCTGTTGCTGATCTGTCACCGTGTGCGCCAGCAGCTGGAGC 486
DB	1151 TGTGTTCTGTCGCTTCTGTTGCTGATCTGTCACCGTGTGCGCCAGCAGCTGGAGC 1210
QY	487 TCGGTCGGGACCTGTGCGCGACCTCAGACACACCGTCACTGATGATCTTTCATCA 546
DB	1211 TCGGTCGGGACCTGTGCGCGACCTCAGACACACCGTCACTGATGATCTTTCATCA 1270
QY	547 CCAGCGTTCTGAGCTTCGGCTTCGGTACGCCAGCGCGCGCGCTTTCAGCGCGACCT 606
DB	1271 CCAGCGTTCTGAGCTTCGGCTTCGGTACGCCAGCGCGCGCGCTTTCAGCGCGACCT 1330
QY	607 ATCTGTGCGCGCTTGGCGCGAGCGGCTGCTCCGACGCGAGCGCGAGTTGCCGAGAG 666
DB	1331 ATCTGTGCGCGCTTGGCGCGAGCGGCTGCTCCGACGCGAGCGCGAGTTGCCGAGAG 1390
QY	667 AACTGGAGCAACTGGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTTTCTCAGCAAAAAG 726
DB	1391 AACTGGAGCAACTGGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTTTCTCAGCAAAAAG 1450
QY	727 AGCTCCGGCGCTGCTGGAGACAGAGGTCACTACCGAGTTTCACGACAGAGCTTCAGG 786
DB	1451 AGCTCCGGCGCTGCTGGAGACAGAGGTCACTACCGAGTTTCACGACAGAGCTTCAGG 1510
QY	787 AGTTCCTTCGGCGACTGTCTTACCTGCTGGAGACGCGCGGCTGCCAGGACCGCGGCT 946

Db	1511	AgTTCTTCGGGCACTGTCTACTCTCTGAGAGACGGGGGTGCCCCAGGACCCGCGCTG	1570
Qy	847	GGGCGTTTGGGACACTCCTCGTGTGGGAGCGCCAGCGCGACAGCCACTTGGTGTCTACCA	906
Db	1571	GGCGGTTTGGGACACTCCTCGTGTGGGAGCGCCAGCGCGACAGCCACTTGGTGTCTACCA	1630
Qy	907	CGCGCTTCCTCTTCGGACTCCTCAGCGCGGAGCGGATCGGACATCGAGCGCCACTTCG	966
Db	1631	CGCGCTTCCTCTTCGGACTCCTCAGCGCGGAGCGGATCGGACATCGAGCGCCACTTCG	1690
Qy	967	GCTGCATGGTTTCAGAGCGTGTGAAGCAGGAGGCCCTCGGTGGGTGCAGGAGCAGGGAC	1026
Db	1691	GCTGCATGGTTTCAGAGCGTGTGAAGCAGGAGGCCCTCGGTGGGTGCAGGAGCAGGGAC	1750
Qy	1027	AGGCGTCCCGGAGTGGCACACAGAGGTGACCGAGGGGCGCAAAAGGCTCGAGAGCACCG	1086
Db	1751	AGGCGTCCCGGAGTGGCACACAGAGGTGACCGAGGGGCGCAAAAGGCTCGAGAGCACCG	1810
Qy	1087	AAGGCCAGAGGAGGAGGAGGAGGAGGCCAACTACCCACTCGAGTTGCTGTACT	1146
Db	1811	AAGGCCAGAGGAGGAGGAGGAGGAGGCCAACTACCCACTCGAGTTGCTGTACT	1870
Qy	1147	GCTGTGTCGAGACGACGAGAGACGCTTTGTGCGCCAGCCCTGTGCCGTTTCCCGGAGC	1206
Db	1871	GCTGTGTCGAGACGACGAGAGACGCTTTGTGCGCCAGCCCTGTGCCGTTTCCCGGAGC	1930
Qy	1207	TGGCGCTGCAGAGTTCGGCTTTCGCGCATGAGACGTGCTGTTCTGAGCTACTGCTGTGA	1266
Db	1931	TGGCGCTGCAGAGTTCGGCTTTCGCGCATGAGACGTGCTGTTCTGAGCTACTGCTGTGA	1990
Qy	1267	GGTGTGCCCTGCTCGACAGGCACTTCGGCTGTATCAGCTGCAGATTTGTTGTCGCGAGG	1326
Db	1991	GGTGTGCCCTGCTCGACAGGCACTTCGGCTGTATCAGCTGCAGATTTGTTGTCGCGAGG	2050
Qy	1327	AGAGAGAGAGAGAGGCTTGGAGAGGGCTCCAGGCGAGCCTGGGTGGCGGCACTCTC	1386
Db	2051	AGAGAGAGAGAGAGGCTTGGAGAGGGCTCCAGGCGAGCCTGGGTGGCGGCACTCTC	2110
Qy	1387	AAGGCACCAAAAACAATGCCAGCCTCCCTTCTTCATCCACTCTTTCAGGCAATGACTG	1446
Db	2111	AAGGCACCAAAAACAATGCCAGCCTCCCTTCTTCATCCACTCTTTCAGGCAATGACTG	2170
Qy	1447	AGCCACTGTGCCATCTGAGCAGCCTCAGCCTGTCGCACTGCAAACTCCCTGACCGGCTC	1506
Db	2171	AGCCACTGTGCCATCTGAGCAGCCTCAGCCTGTCGCACTGCAAACTCCCTGACCGGCTC	2230
Qy	1507	GCCGAGACCTTTCTGAGGCCCTCAGGGCAGCCCCCGCACTGACGAGCTGGCGCTCTCC	1566
Db	2231	GCCGAGACCTTTCTGAGGCCCTCAGGGCAGCCCCCGCACTGACGAGCTGGCGCTCTCC	2290
Qy	1567	ACAACAGCCTCAGTGAAGCGGACTGGGTATGCTGAGTGAGGCGCTACGCTTGGCGGAGT	1626
Db	2291	ACAACAGCCTCAGTGAAGCGGACTGGGTATGCTGAGTGAGGCGCTACGCTTGGCGGAGT	2350
Qy	1627	GCAGGTCACAGCGTTCAGGGTACAGTTCCTGAGACCCCGAGCGGGTCTCCAGTACCTCG	1686
Db	2351	GCAGGTCACAGCGTTCAGGGTACAGTTCCTGAGACCCCGAGCGGGTCTCCAGTACCTCG	2410
Qy	1687	TGGGTATGCTTTGGCAGAGCCCGCCTGACCAACCTGGATCTCAGCGGCTGCCAATGTC	1746
Db	2411	TGGGTATGCTTTGGCAGAGCCCGCCTGACCAACCTGGATCTCAGCGGCTGCCAATGTC	2470
Qy	1747	CGCCCCCATCGTGACTTCTGTGTCAGTCTCTGCAGCACCGGATTCGCGCTCTGCAGA	1806
Db	2471	CGCCCCCATCGTGACTTCTGTGTGTCAGTCTCTGCAGCACCGGATTCGCGCTCTGCAGA	2530
Qy	1807	CCCTCAGTCTGGCCTCTGTGGAGCTGAGCGGAGCAGTCACTACAGGAGCTTCAGGCTGTGA	1866
Db	2531	CCCTCAGTCTGGCCTCTGTGTGAGCTGAGCGGAGCAGTCACTACAGGAGCTTCAGGCTGTGA	2590
Qy	1867	AGAGAGCAAGCCGGATCTCGTTCATCACACCCAGCGCTTGAAGCGGCCACCCCAAACTTC	1926

DB 2521 ACA3GCGTCCAGTCTCTAGATCTTTGGGAAGGAGACTAGCGCAGGTGGAGACACGCA 2580
QY 2507 GAACCCCGTCTGGTGGGAGACATGACACATGCTGGTGGAGAGCCGCCATGCACTG 2566
DB 2581 GAACCCCGTCTGGTGGGAGACATGACACATGCTGGTGGAGAGCCGCCATGCACTG 2640
QY 2567 ACGTAAATCCCTGTGGACTCATTTCTGTTCTTATTACACTGGCCAGGCGTG 2626
DB 2641 ACGTAAATCCCTGTGGACTCATTTCTGTTCTTATTACACTGGCCAGGCGTG 2700
QY 2627 TACAATACAGTGGTGTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 2686
DB 2701 TACAATACAGTGGTGTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 2760
QY 2687 AAA 2689
DB 2761 AAA 2763
RESULT 4
ABS63487
ID ABS63487 standard; cDNA; 2054 BP.
AC
XX ABS63487;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human leucine-rich repeat small intestine 1 (HLRSL1), EST #2.
XX
KW Human; human leucine-rich repeat small intestine 1; HLRSL1; asthma;
KW proliferative disorder; gastrointestinal disorder; renal disorder;
KW neural disorder; reproductive disorder; calcium regulation; apoptosis;
KW immune system; anaemia; human immune deficiency virus; HIV; cancer;
KW blood coagulation disorder; autoimmune disorder; allergic reaction;
KW inflammatory condition; cardiovascular disorder; ischaemia;
KW neurological disorder; infectious disease; cytokine production;
KW expressed sequence tag; EST; gene; ss.
XX
OS Homo sapiens.
XX
XX
PN WO200261096-A2.
XX
XX 08-AUG-2002.
XX
PF 20-DEC-2000; 2000US-257774P.
XX
XX 22-DEC-2000; 2000US-257774P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Feder J, Ramanathan C, Mintier G;
XX
XX WPI; 2002-6:9252/66.
XX
XX New isolated nucleic acid molecules encoding HLRSL1 polypeptides, or
XX their fragments and homologues, useful for preventing, treating and
XX ameliorating medical conditions, e.g. proliferative, gastrointestinal,
XX or renal disorders -
XX
PS Example 1; Page 222-224; 336pp; English.
XX
XX The invention relates to isolated nucleic acid molecules (I) encoding
XX human leucine-rich repeat small intestine 1 (HLRSL1) polypeptides.
XX The nucleic acid molecules and polypeptides are useful for preventing,
XX treating and ameliorating medical conditions, such as proliferative,
XX gastrointestinal, renal, neural, or reproductive disorders; or disorders
XX related to aberrant calcium regulation or apoptosis modulation, either
XX directly or indirectly. They are also useful for treating, preventing
XX and/or diagnosing diseases, disorders and/or conditions of: immune system
XX by activating or inhibiting the proliferation, differentiation, or
XX mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
XX anaemia; immunologic deficiency syndromes, e.g. human immune deficiency

CC virus (HIV) infection, HTLV-BLV infection; blood coagulation disorders,
CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
CC myasthenia gravis; asthma or allergic reactions; inflammatory
CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
CC ischaemia, aneurysms; neurological disorders, e.g. Alzheimer's disease,
CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
CC pneumonia, or viral, bacterial, and fungal infections. The HLRSL1
CC polypeptides are useful for modulating cytokine production, antigen
CC presentation, or other processes such as boosting immune responses.
CC ABS63485-ABS63504 represent HLRSL1 coding sequences and PCR primers of
XX the invention.
SQ Sequence 2054 BP; 473 A; 631 C; 607 G; 343 T; 0 other;
Query Match 75.1%; Score 2020.2; DB 24; Length 2054;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 665 GGAACTGGAGCAACTGGAGCTTCGTGGCTCCAAAGTCAGACGCTGTCTCTCAGCAAAA 724
DB 1 GGAACTGGAGCAACTGGAGCTTCGTGGCTCCAAAGTCAGACGCTGTCTCTCAGCAAAA 60
QY 725 GGAGCTGCGGCGCTGTGGAGACAGAGTCACTTCCAGTTCATCGACAGAGTTCCA 784
DB 61 GGAGCTGCGGCGCTGTGGAGACAGAGTCACTTCCAGTTCATCGACAGAGTTCCA 120
QY 785 GGAGTTCCTCGGCGCACTGTCTTACCTCTGGAGACGCGGCGGTCCAGACCGCGCG 844
DB 12 GGAGTTCCTCGGCGCACTGTCTTACCTCTGGAGACGCGGCGGTCCAGACCGCGCG 180
QY 845 TGGCGGCGTGGGACACTCTCTGCTGGGAGCCCGCAGCCGACAGCCACTTGGTCTCAC 904
DB 18 TGGCGGCGTGGGACACTCTCTGCTGGGAGCCCGCAGCCGACAGCCACTTGGTCTCAC 240
QY 905 CACGCGCTTCTCTTTCGAGCTGTGAGCGGAGCGGATGCGCGACATCGAGCGCACTT 964
DB 241 CACGCGCTTCTCTTTCGAGCTGTGAGCGGAGCGGATGCGCGACATCGAGCGCACTT 300
QY 965 CGGCTGCATGCTTTCAGAGCGGTGTGAACGAGAGCGCTTGGGTGGTGCAGGGACAGG 1024
DB 301 CGGCTGCATGCTTTCAGAGCGGTGTGAACGAGAGCGCTTGGGTGGTGCAGGGACAGG 360
QY 1025 ACAGGCTGCCCGGAGTGGACACAGAGTGTACCGAGGGGCGCAAGGCTCGAGGACAC 1084
DB 361 ACAGGCTGCCCGGAGTGGACACAGAGTGTACCGAGGGGCGCAAGGCTCGAGGACAC 420
QY 1085 CGAAGCGCAGAGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1144
DB 421 CGAAGCGCAGAGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 1145 CTGCTGTACGAGCGCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1204
DB 481 CTGCTGTACGAGCGCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 1205 GCTGCGCTGACGAGGAGTGGCTTCTGCGGATGACGCTGCTGTCTGAGTACTGGCT 1264
DB 541 GCTGCGCTGACGAGGAGTGGCTTCTGCGGATGACGCTGCTGTCTGAGTACTGGCT 600
QY 1265 GAGGTGCTGCCCTGTGACAGGCACTGCGGCTGATCAGCTGCAGATTGGTTCGCGCA 1324
DB 601 GAGGTGCTGCCCTGTGACAGGCACTGCGGCTGATCAGCTGCAGATTGGTTCGCGCA 660
QY 1325 CGAG 1384
DB 661 CGAG 720
QY 1385 TCAAGGACCAACAAACAACTGCCAGCTCCCTTCTTATCCACTCTTTCAGGCAATGAC 1444
DB 721 TCAAGGACCAACAAACAACTGCCAGCTCCCTTCTTATCCACTCTTTCAGGCAATGAC 780
QY 1445 TGACCCACTGTGCCATCTGAGAGAGCTTCAAGCTGTCACAGTGCACAACTCCCTGACGCGT 1504

Db 781 TGACCCACCTGTGGCATCTGACACCCCTCAGCTGTGCCACTGCAAACTCCCTGACCGGT 840
Qy 1505 CTGGCGAGACCTTTCTGAGGCTCTGAGGCGCTGAGGCGAGCCCGGCACTGACGAGCTGGGCTCTCT 1564
Db 841 CTGGCGAGACCTTTCTGAGGCTCTGAGGCGCTGAGGCGAGCCCGGCACTGACGAGCTGGGCTCTCT 900
Qy 1565 CCACACAGGCTCAGTGAAGCGGAGCTGCGTATGCTGAGTGAGGCTGAGCTGAGCGCGCA 1624
Db 901 CCACACAGGCTCAGTGAAGCGGAGCTGCGTATGCTGAGTGAGGCTGAGCTGAGCGCGCA 960
Qy 1625 GTGAGGCTGACAGCGGTACAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 1684
Db 961 GTGAGGCTGACAGCGGTACAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 1020
Qy 1685 GGTGGTATGCTTGGCGAGACCCCGGCTGACACCCCTGATCTGAGCGCTGAGCGCT 1744
Db 1021 GGTGGTATGCTTGGCGAGACCCCGGCTGACACCCCTGATCTGAGCGCTGAGCGCT 1080
Qy 1745 GCGCCCGCCCATGCTGACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1804
Db 1081 GCGCCCGCCCATGCTGACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy 1805 GACCTCAGTCTGGGCTCTGTGAGCTGAGGCGAGCTGACTACAGAGCTTCAGGCTGT 1864
Db 1141 GACCTCAGTCTGGGCTCTGTGAGCTGAGGCGAGCTGACTACAGAGCTTCAGGCTGT 1200
Qy 1865 GAAGAGACCAAGCCGATCTGGTCACTCACACCCAGCGCTGACCGCCACCCACCAACC 1924
Db 1201 GAAGAGACCAAGCCGATCTGGTCACTCACACCCAGCGCTGACCGCCACCCACCAACC 1260
Qy 1925 TCCCAAGAACCTACTCTCGACTCTGAGGCTCTGGTGGCGAGGAGGCTGAGAACCC 1984
Db 1261 TCCCAAGAACCTACTCTCGACTCTGAGGCTCTGGTGGCGAGGAGGCTGAGAACCC 1320
Qy 1985 TAGTCAAGTCCCTGTGAGAGAGAGCGCCCTTCCAGGCGAGGAGCACTTCTCTGG 2044
Db 1321 TAGTCAAGTCCCTGTGAGAGAGAGCGCCCTTCCAGGCGAGGAGCACTTCTCTGG 1380
Qy 2045 CTTTGGGAACTTTGAGTGGAGAGGCGCGAGAGCGATGTGGAGGCGCCAGACAGG 2104
Db 1381 CTTTGGGAACTTTGAGTGGAGAGGCGCGAGAGCGATGTGGAGGCGCCAGACAGG 1440
Qy 2105 CACCTGCGCGCTGAGGAGAGGCGCGAGGCTGCGCTCTCTCCACCTGGGTAACC 2164
Db 1441 CACCTGCGCGCTGAGGAGAGGCGCGAGGCTGCGCTCTCTCCACCTGGGTAACC 1500
Qy 2165 CTTCTCCCGAGCCCACTACTCTCAACCACTTCTCTCTGTGAGAGCTCCAGCAT 2224
Db 1501 CTTCTCCCGAGCCCACTACTCTCAACCACTTCTCTCTGTGAGAGCTCCAGCAT 1560
Qy 2225 TCCCTTTGAAACACCCCGCCAGCCCAAGCCCAATTAATCAGAGGAGAGCTCCCAATTAA 2284
Db 1561 TCCCTTTGAAACACCCCGCCAGCCCAAGCCCAATTAATCAGAGGAGAGCTCCCAATTAA 1620
Qy 2285 CTAAGCAGCTTACCTGGGCGAGATAACCTTCTACTGCTGATCCCACTGAGGCTGAGG 2344
Db 1621 CTAAGCAGCTTACCTGGGCGAGATAACCTTCTACTGCTGATCCCACTGAGGCTGAGG 1680
Qy 2345 CCCAAGCAGCCCAAGCTTATGCCACATAGACTGAGGAGTGGAGGCTGAGGCTGAGGCTGAGG 2404
Db 1681 CCCAAGCAGCCCAAGCTTATGCCACATAGACTGAGGAGTGGAGGCTGAGGCTGAGGCTGAGG 1740
Qy 2405 CTGTTAGGAATGAGACCTTCCCTGAGGCTATGGCCAGGCGCCAGAGCGTCCAGTGTCTG 2464
Db 1741 CTGTTAGGAATGAGACCTTCCCTGAGGCTATGGCCAGGCGCCAGAGCGTCCAGTGTCTG 1800
Qy 2465 AGATCTTTGGGAGGAGCTAGGCGAGGCTGAGGAGGCGAGGAGGAGGAGGAGGAGGAGGAGG 2524
Db 1801 AGATCTTTGGGAGGAGCTAGGCGAGGCTGAGGAGGCGAGGAGGAGGAGGAGGAGGAGGAGG 1860
Qy 2525 GGAAGCATGACCATGCTGGTGGTGGAGGAGGCGCCCATGCACTGAGCGTAAATTCCTCTGTG 2584
Db 1861 GGAAGCATGACCATGCTGGTGGTGGAGGAGGCGCCCATGCACTGAGCGTAAATTCCTCTGTG 1920

Qy 2585 GACTCATTCTGTGTTCTTATTACCTGCGCAGCGCTGTACAAATACAGGTCGGTGC 2644
Db 1921 GACTCATTCTGTGTTCTTATTACCTGCGCAGCGCTGTACAAATACAGGTCGGTGC 1980
Qy 2645 TCACAAAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2689
Db 1981 TCACAAAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2025

RESULT 5

ABX97042
ID ABX97042 standard; cDNA; 2109 BP.

XX ABX97042;

XX 20-MAY-2003 (first entry)

XX Human NOV18a cDNA.

XX NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
XX hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
XX human; gene; ss.

XX Homo sapiens.

XX WO200272757-A2.

XX 19-SEP-2002.

XX 08-MAR-2002; 2002WO-US06908.

XX 08-MAR-2001; 2001US-274101P.

XX 08-MAR-2001; 2001US-274194P.

XX 08-MAR-2001; 2001US-274281P.

XX 08-MAR-2001; 2001US-274322P.

XX 09-MAR-2001; 2001US-274849P.

XX 12-MAR-2001; 2001US-275235P.

XX 13-MAR-2001; 2001US-275578P.

XX 13-MAR-2001; 2001US-275579P.

XX 13-MAR-2001; 2001US-275601P.

XX 14-MAR-2001; 2001US-276000P.

XX 16-MAR-2001; 2001US-276776P.

XX 19-MAR-2001; 2001US-276894P.

XX 20-MAR-2001; 2001US-277239P.

XX 20-MAR-2001; 2001US-277321P.

XX 20-MAR-2001; 2001US-277327P.

XX 21-MAR-2001; 2001US-277791P.

XX 22-MAR-2001; 2001US-277833P.

XX 23-MAR-2001; 2001US-278152P.

XX 26-MAR-2001; 2001US-278894P.

XX 27-MAR-2001; 2001US-278999P.

XX 27-MAR-2001; 2001US-279366P.

XX 28-MAR-2001; 2001US-279344P.

XX 30-MAR-2001; 2001US-277338P.

XX 30-MAR-2001; 2001US-279959P.

XX 30-MAR-2001; 2001US-280233P.

XX 02-APR-2001; 2001US-280802P.

XX 02-APR-2001; 2001US-280822P.

XX 02-APR-2001; 2001US-280822P.

XX 04-APR-2001; 2001US-281194P.

XX 04-APR-2001; 2001US-283675P.

XX 30-MAY-2001; 2001US-287424P.

XX 03-MAY-2001; 2001US-288066P.

XX 03-MAY-2001; 2001US-288342P.

XX 03-MAY-2001; 2001US-288528P.

XX 15-MAY-2001; 2001US-291190P.

XX 16-MAY-2001; 2001US-291099P.

XX 16-MAY-2001; 2001US-291240P.

XX 30-MAY-2001; 2001US-294485P.

XX 31-MAY-2001; 2001US-294889P.

XX 31-MAY-2001; 2001US-294899P.

XX 18-JUN-2001; 2001US-299027P.


```

Db 1389 GGTGCTGCGCTGCTGCACAGC3CACTGCGGCTGATCAGCTCAGATTGGTTGCTGCGCAGG 1446
QY 1327 AGAGAGAGAGAGAGAGCTGCGGAAACGCGCTCAGGCGCAGCTGGTGGCGGAGTTCTC 1386
XX 1449 AGAGAGAGAGAGAGAGCTGCGGAAACGCGCTCAGGCGCAGCTG-----1492
Db 1387 AAGGACACACAAACAACTGCCAGCCTCCCTTTTCATCCACTCTTTCAGGCAATGACTG 1446
QY 1493 --GGCACACAAACAACTGCCAGCCTCCCTTTTCATCCACTCTTTCAGGCAATGACTG 1550
Db 1447 ACCCATGTGCAATCTGAGCAGCCTCAGCTGTGCCACTGCAAACTCCCTTGAGCGGGTCT 1506
QY 1551 ACCCATGTGCAATCTGAGCAGCCTCAGCTGTGCCACTGCAAACTCCCTTGAGCGGGTCT 1610
Db 1507 GCCGAGACCTTTCTCAGGCGCTGAGGCGAGCCCGGACCTGACGAGCTGGGCTCTCTCC 1566
QY 1611 GCCGAGACCTTTCTCAGGCGCTGAGGCGAGCCCGGACCTGACGAGCTGGGCTCTCTCC 1670
Db 1567 ACAACAGGCTCAGTGTGAGCGGAGCTGCGTATGCTGAGTGAGGGCTTAGCGCGCAGT 1626
QY 1671 ACAACAGGCTCAGTGTGAGCGGAGCTGCGTATGCTGAGTGAGGGCTTAGCGCGCAGT 1730
Db 1627 GCAGGCTGCAGAGCTGAGGCTACAGCTGCTGACCCCGGAGGCTCCAGTACTG 1686
QY 1731 GCAGGCTGCAGAGCTGAGGCTACAGCTGCTGACCCCGGAGGCTCCAGTACTG 1790
Db 1687 TGGGTATGCTTGGCAGAGTCCCGCTGACACCCCTGATCTCAGCGGCTGCCAATGCT 1746
QY 1791 TGGGTATGCTTGGCAGAGTCCCGCTGACACCCCTGATCTCAGCGGCTGCCAATGCT 1850
Db 1747 CCGCCCGCATGTGACCTAATCTGTGTGAGTCTGACAGCCAGGATGGGCTGCAGA 1806
QY 1851 CCGCCCGCATGTGACCTAATCTGTGTGAGTCTGACAGCCAGGATGGGCTGCAGA 1910
Db 1807 CCTCAGTGTGCTCTGTGAGCTGAGCGAGCTCCTACAGAGCTTCAAGCTGTGA 1866
QY 1911 CCTCAGTGTGCTCTGTGAGCTGAGCGAGCTCCTACAGAGCTTCAAGCTGTGA 1970
Db 1867 AGAGAGAGAGAGAGAGCTGCTGCTACACACCCAGCGCTGGAGGCGACCCAGCTC 1926
QY 1971 AGAGAGAGAGAGAGAGCTGCTGCTACACACCCAGCGCTGGAGGCGACCCAGCTC 2030
Db 1927 CCAAGGAACTCATCTCGAATCTTCTGAGCTCTGTGCGCAGAGCAGGCTGGAAAGACCTA 1986
QY 2031 CCAAGGAACTCATCTCGAATCTTCTGAGCTCTGTGCGCAGAGCAGGCTGGAAAGACCTA 2090
Db 1987 GTCAAGTCCCTGTGAG 2005
QY 2091 GTCAAGTCCCTGTGAG 2109

```

RESULT 6

AAL47132

ID AAL47132 standard; DNA; 5556 BP.

XX AC AAL47132;

XX DT 23-AUG-2002 (first entry)

XX DE Pyrin domain containing protein NALP6/pv9 coding sequence.

XX KW Pyrin domain; PYD domain; anti-inflammatory; antiparkinsonian;

XX KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;

XX KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;

XX KW rephrotropic; osteopathic; nootropic; intracellular signal transduction;

XX KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;

XX KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;

XX KW osteoarthritis; glomerulonephritis; gene; ds.

XX OS Unidentified.

XX PN W020240668-A2.

XX

```

PJ 23-YAY-2002.
XX 30-OCT-2001; 2001WO-EPI2545.
XX 15-NOV-2003; 2000DE-1056687.
XX 30-NOV-2000; 2000DE-1059595.
XX (APOT-) APOTEC RES & DEV LTD.
XX Tschopp J, Martinon F;
XX WP; 2002-427093/45.
XX P-PSDB; AAO17860.
XX New DNA encoding protein with pyrin domain, useful for treating
XX diseases involving impaired signal transduction, particularly
XX inflammation, also proteins and antibodies -
XX Claim 5; Fig 1; 116pp; German.
XX The present invention relates the DNA and their encoded proteins, where
XX the proteins contain at least one PYD (pyrin) domain. These can be used
XX to treat diseases associated with impaired intracellular signal
XX transduction, particularly inflammation such as psoriasis,
XX arteriosclerosis, bacterial or viral infections (particularly meningitis
XX and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX and Parkinson's diseases. The present sequence is a coding sequence of
XX the invention.
XX Sequence 5556 BP; 914 A; 1856 C; 1879 G; 907 T; 0 other;
XX
XX Query Match 58.4%; Score 1571.2; DB 24; Length 5556;
XX Best Local Similarity 89.6%; Pred. No. 5e-272;
XX Matches 1772; Conservative 0; Mismatches 38; Indels 168; Gaps 1;
QY 7 CGTGGGCGCAGCCTGGCTGACCTCATCTGACCCAGTGCCTGACCGCGCGCGCGCGG 66
Db 731 CGGCGACCGCAGCCTGGCTGACCTCATCTGACCGCGCGCGCGCGCGCGCGG 790
QY 67 TGCGCAGATGCTGGCGCCAGCCAGCGCTGCTTTCATCTGAGCGCGCGCGCGG 126
Db 791 TGCGCAGATGCTGGCGCCAGCCAGCGCTGCTTTCATCTGAGCGCGCGCGG 850
QY 127 TGCGCGGCTGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 186
Db 851 TGCGCGGCTGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 910
QY 187 GCGCGGCTGTAGCGGCTGTGATGAGTAAGCGCTGCTGCCACGCGCTCTCTGCTGG 246
Db 911 GCGCGGCTGTAGCGGCTGTGATGAGTAAGCGCTGCTGCCACGCGCTCTCTGCTGG 970
QY 247 TGACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 306
Db 971 TGACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 1030
QY 307 CCAGAGTGTGCGCGCTTCTCCGACAAAGCAAGAGATTTCTACAAGTTCTTCCCGG 366
Db 1031 CCAGAGTGTGCGCGCTTCTCCGACAAAGCAAGAGATTTCTACAAGTTCTTCCCGG 1090
QY 367 ATGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 426
Db 1091 ATGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 1150
QY 427 TGTGCTTGTGCGCTTCTGCTGTGCTGATGCTGTGACCGCTGCTGCCACGCGCTGAGC 486
Db 1151 TGTGCTTGTGCGCTTCTGCTGTGCTGATGCTGTGACCGCTGCTGCCACGCGCTGAGC 1210
QY 487 TCGGTGCGGACCTGTGCGCGCACTCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 546
Db 1211 TCGGTGCGGACCTGTGCGCGCACTCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1270
QY 547 CCAGCGTCTGAGCTCGGCTCCGCTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606

```



```
Db 1271 CCAGGCTTCTGAGCTCGGCTCCGGTAGCCGAGCGGCCCGGTTGACGGCGACCTGCGCA 1330
Qy 607 ATCTGTCGGGCTGCGCGCGAGGCGGTCCTCGGACGAGGCGGCGAGTTTCCGAGAGG 666
Db 1331 ATCTGTCGGGCTGCGCGCGAGGCGGTCCTCGGACGAGGCGGCGAGTTTCCGAGAGG 1389
Qy 667 AACTGGAGCACTGGAGCTTCTGCTGCTCCAAAGTGCAGACGCTGTTCTTCAGCAAAAGG 726
Db 1391 AACTGGAGCACTGGAGCTTCTGCTGCTCCAAAGTGCAGACGCTGTTCTTCAGCAAAAGG 1450
Qy 727 AGCTCCCGGCGCTGCTGGAGACAGAGGTCACCTACCAAGTTCAATGACAGAGCTTCCAGG 786
Db 1451 AGCTCCCGGCGCTGCTGGAGACAGAGGTCACCTACCAAGTTCAATGACAGAGCTTCCAGG 1510
Qy 787 AGTTCTTCCGGCACTGCTTCTTCTTCTGCTGGAGGAGCGGGGTCGCGAGGACCGGCTG 846
Db 1511 AGTTCTTCCGGCACTGCTTCTTCTTCTGCTGGAGGAGCGGGGTCGCGAGGACCGGCTG 1570
Qy 847 GCGGCTTGGGACACTCTCTGCTGGGAGCGCCAGCGCACAGGCACTTGGTGTCCACCA 906
Db 1571 GCGGCTTGGGACACTCTCTGCTGGGAGCGCCAGCGCACAGGCACTTGGTGTCCACCA 1630
Qy 907 CCGCTTCTCTTCCGACTGCTGAGCGGAGCGGATCGGACATCGAGCGCACTTTCG 966
Db 1631 CCGCTTCTCTTCCGACTGCTGAGCGGAGCGGATCGGACATCGAGCGCACTTTCG 1690
Qy 967 GCTGATGTTTTCAGAGGCTGTAAGAGGAGGCGCTCGGGTGGGTGCAGGACAGGAGAC 1026
Db 1691 GCTGATGTTTTCAGAGGCTGTAAGAGGAGGCGCTCGGGTGGGTGCAGGAGGAGAC 1750
Qy 1027 AGGGCTGCCCGGAGTGGSCACAGAGGTGACAGGAGGCGCCAAAGGCTCGAGGACACCG 1086
Db 1751 AGGGCTGCCCGGAGTGGSCACAGAGGTGACAGGAGGCGCCAAAGGCTCGAGGACACCG 1810
Qy 1087 AAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1146
Db 1811 AAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1870
Qy 1147 GCTGTAAGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1206
Db 1871 GCTGTAAGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1930
Qy 1207 TGGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1266
Db 1931 TGGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1990
Qy 1267 GGTGCTGCGCTGCTGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1326
Db 1991 GGTGCTGCGCTGCTGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2050
Qy 1327 AGAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1386
Db 2051 AGAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2110
Qy 1387 AAGGACCAACAAACAACTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1433
Db 2111 TGGGACCAACAACTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2170
Qy 1434 ----- 1432
Db 2171 CACCTCCACAGACCTCGGCTCCTCCAGGGAAGGCTTTTCCAGAGGTTCTTTGAATA 2230
Qy 1434 ----- 1433
Db 2231 TAGCTCCAAATTCAGCCCTGSCCAGAGGGCTTGGCATCTGTTGAGAGGATGAATGTCAGG 2290
Qy 1434 -----CAGGCAATGACTGACCCACTGTGCC 1458
Db 2291 TGTGGCAGGGCTGGGCTGGGACCCAAAGACCCATGCAATGACTGACCCACTGTGCC 2350
Qy 1459 ATCTGAGAGCCTCAGCTGTCCCACTGCAAACTCCCTGACGCGGTGTCGCCGAGACCTTT 1518
```

```
2351 ATCTGAGCAGCCTCAGCGTGTCCACTGCAAACTCCCTGACGCGGTCTGCCGAGACCTTT 2410
Qy 1519 CTGAGGCCCTGAGGCGAGCGCCCGCACTGACGAGCTGGGCTCTCTCCACAAAGGCTCA 1578
Db 2411 CTGAGGCCCTGAGGCGAGCGCCCGCACTGACGAGCTGGGCTCTCTCCACAAAGGCTCA 2470
Qy 1579 GTGAGCGGAGCTGCGTATGCTGAGTGAGGGCTAGCCTTGGCCGAGTGCGAGGTCGAGA 1638
Db 2471 GTGAGCGGAGCTGCGTATGCTGAGTGAGGGCTAGCCTTGGCCGAGTGCGAGGTCGAGA 2530
Qy 1639 CGGTGAGGGTACAGCTGCTGACCCCGAGGAGGCTCGAGTACCTGCTGGTGGTATGCTTC 1698
Db 2531 CGGTGAGGGTACAGCTGCTGACCCCGAGGAGGCTCGAGTACCTGCTGGTGGTATGCTTC 2590
Qy 1699 GGAGAGCGCCGCGCTTGAACACCTCGATCTCAGCGGCTGCCAACTGCCCGCCCATGG 1758
Db 2591 GGAGAGCGCCGCGCTTGAACACCTCGATCTCAGCGGCTGCCAACTGCCCGCCCATGG 2650
Qy 1759 TGACCTACCTGTGTGAGTCTCTGCAGCACCAGGAGTGGGCTTGCAGACCTCAGTCT 1816
Db 2651 TGACCTACCTGTGTGAGTCTCTGCAGCACCAGGAGTGGGCTTGCAGACCTCAGCT 2708

RESULT 7
AAS00033
CD AAS00033 standard; cDNA; 6461 BP.
XX
AC AAS00033;
XX
DT 10-MAY-2001 (first entry)
XX
DE Human ATLAS-2-encoding cDNA.
XX
KW Human; Activated T-lymphocyte associated sequence 2; ATLAS-2; antibody;
cytokine receptor; autoimmune disorder; immune disorder; cancer;
T-lymphocyte-associated disorder; cell-proliferation disorder; tumour;
cell differentiation disorder; immune deficiency disorder; malignancy;
viral infection; bacterial infection; fungal infection; metabolism;
chromosome 1p15.5; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..5556
FT /*tag= a
FT /product= "ATLAS-2"
XX
PN WO200114564-A2.
XX
PD 01-MAR-2001.
XX
PF 18-AUG-2000; 2000WO-US22699.
XX
PR 20-AUG-1999; 99US-0150105.
PR 28-APR-2000; 2000US-0560948.
PR 28-APR-2000; 2000US-0560948.
PR 28-APR-2000; 2000US-0561533.
XX
PA (CURA-) CURAGEN CORP.
PA (BIO-) BIOGEN INC.
XX
PI Peyman JA, Green CD, Hsu A, Browning JA, Carulli J;
XX
PL WPI; 2001-218453/22.
DR P-PSDB; AAU00023.
XX
PT New isolated activated T lymphocyte associated sequences for treating
or preventing immune system associated disorders such as autoimmune
disorder, immune disorder, and T-lymphocyte-associated disorder -
XX
PS Claim 6; Fig 2; 114pp; English.
XX
```


CC The sequence encodes human activated T-lymphocyte associated sequence 2,
CC ATLAS-2. ATLAS-2 is related by homology to cytokine receptors and its
CC gene is located on chromosome 11p15.5. ATLAS proteins, polynucleotides
CC and antibodies are useful for treating/preventing conditions associated
CC with an autoimmune disorder, immune disorder, T-lymphocyte-associated
CC disorder, cell-proliferation disorder, cell differentiation disorder,
CC and immune deficiency disorder and for screening for a modulator of
CC activity or of latency or predisposition to an immune disorder.
CC ATLAS proteins, polynucleotides and antibodies are useful in therapeutic
CC or prophylactic treatment of diseases associated with cell-proliferation
CC (e.g. cancers, malignancies and tumours). The polynucleotides are useful
CC in gene therapy, to detect ATLAS mRNA or a genetic lesion in an ATLAS
CC gene, to modulate ATLAS activity, to screen drugs or compounds that
CC modulate ATLAS activity or expression and to treat disorders
CC characterised by insufficient or excessive production of ATLAS protein or
CC production of ATLAS protein forms that have decreased or aberrant
CC activity compared to ATLAS wild type protein and in tissue typing to
CC identify individuals. The antibodies are useful for localisation/
CC quantitation, isolation and detection of ATLAS and to monitor protein
CC levels in tissue. ATLAS is useful for treating/preventing infection by
CC bacteria, viruses and fungi, affecting bodily characteristics, e.g.
CC biohythms, fertility or metabolism, affecting behavioural
CC characteristics, and for providing analgesic effects. A host cell
CC containing the polynucleotide is useful to produce non-human transgenic
CC animals.
XX
XX
SQ Sequence 6461 BP; 1079 A; 2116 C; 2150 G; 1086 T; 0 other;

Query Match 58.4%; Score 1571.2; DB 22; Length 6461;
Best Local Similarity 89.6%; Pred. No. 5.1e-272;
Matches 1772; Conservative 3; Mismatches 38; Indels 168; Gaps 1;

QY 7 CQTGGGCGGAGCTGCTGACTGATCTCTGGACCAAGTCCCGACCGCGGCGCGCG 66
DB 731 CGGCGACGCGAGCTGCTGACTGATCTCTGGACCAAGTCCCGACCGCGGCGCGCG 790
QY 67 TCGCGCAGATGTCGCCCGACGCGGAGCGCTCTTCATCTCGAGCGCGCGAGCG 126
DB 791 TCGCGCAGATGTCGCCCGACGCGGAGCGCTCTTCATCTCGAGCGCGCGAGCG 850
QY 127 TCGCGCGCTCGGGGGCGCGGAGCGCGCTCTTCAGAGAGCGCGCGCGAGCG 186
DB 851 TCGCGCGCTCGGGGGCGCGGAGCGCGCTCTTCAGAGAGCGCGCGCGAGCG 910
QY 187 GCGCGCGCTTAGCGCGCTGCTGAGTGAAGCGCTCTTCATCTCGAGCGCGCGAGCG 246
DB 911 GCGCGCGCTTAGCGCGCTGCTGAGTGAAGCGCTCTTCATCTCGAGCGCGCGAGCG 970
QY 247 TGACCAACGCGCGCGCGCGCGCGCGCGCGCTCTGTTCCCGCGAGTGG 306
DB 971 TGACCAACGCGCGCGCGCGCGCGCGCGCGCTCTGTTCCCGCGAGTGG 1030
QY 307 CCGAGGTGGGGGCTTCTCGACAGAGAGAGAGAGTATTTCTAGAGTCTTCGGG 366
DB 1031 CCGAGGTGGGGGCTTCTCGACAGAGAGAGAGAGTATTTCTAGAGTCTTCGGG 1090
QY 367 ATGAG 426
DB 1091 ATGAG 1150
QY 427 TGTGTTTGTGCGCTTGTGCTGGATGCTGTGACCGTCTCGCGAGAGAGAGAGAGAG 486
DB 1151 TGTGTTTGTGCGCTTGTGCTGGATGCTGTGACCGTCTCGCGAGAGAGAGAGAGAG 1210
QY 497 TCGGTTGGGAG 546
DB 1211 TCGGTTGGGAG 1270
QY 547 CCAGGCTTCTGAGCTCGGCTCGGCTAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCA 606
DB 1271 CCAGGCTTCTGAGCTCGGCTCGGCTAGCGAGCGCGCGCGCGCGCGCGCGCGCGCA 1330
QY 607 ATCTGTGCGCGCTGGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666

DB 1331 ATCTGTGCGCGCTGGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1390
QY 667 AACTGGAGAGAACTGGAGCTTGGCTTCCAAAGTGGAGAGAGAGAGAGAGAGAGAGAG 726
DB 1391 AACTGGAGAGAACTGGAGCTTGGCTTCCAAAGTGGAGAGAGAGAGAGAGAGAGAGAG 1450
QY 727 AGCTGCCGCGCGCTGGAG 786
DB 1451 AGCTGCCGCGCGCTGGAG 1510
QY 787 AGTTCTCTGGCGCACTGTCTCTAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
DB 1511 AGTTCTCTGGCGCACTGTCTCTAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1570
QY 847 GCGCGCTTGGAG 906
DB 1571 GCGCGCTTGGAG 1630
QY 907 GCGCGCTTCTCTTGGAG 966
DB 1631 GCGCGCTTCTCTTGGAG 1690
QY 967 GCTGCATGCTTTCAG 1026
DB 1691 GCTGCATGCTTTCAG 1750
QY 1027 AGGCTGTCGCCCGAG 1086
DB 1751 AGGCTGTCGCCCGAG 1810
QY 1087 AAG 1146
DB 1811 AAG 1870
QY 1147 GCCTGTACAG 1206
DB 1871 GCCTGTACAG 1930
QY 1207 TGGCGCTGAG 1266
DB 1931 TGGCGCTGAG 1990
QY 1267 GGTCTGCTGCTGAG 1326
DB 1991 GGTCTGCTGCTGAG 2050
QY 1327 AGAAG 1386
DB 2051 AGAAG 2110
QY 1387 AAGGAG 1433
DB 2111 TGGGAG 1470
QY 1434 ----- 1433
DB 2171 CACCTTCACAG 2230
QY 1434 ----- 1433
DB 2231 TAGCTCCAAATTCAGCGCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2290
QY 1434 ----- 1433
DB 2291 TGTGGAG 2350
QY 1459 ATCTGAG 1518
DB 2351 ATCTGAG 2410
QY 1519 CTGAGGAG 1578

Db 2411 CTAGAGCCCTGAGGGAGCCCGCCACATGACGAGCTGGCCCTCTCCACACAGGCTCA 2470
 QY 1579 GTGAGCGGAGCTGCGTATGCTGAGTGAGGCGCTAGCCTGGCGCGAGTGCAGGTGCAGA 1638
 Db 2471 GTGAGCGAGAGCTGCGTATGCTGAGTGAGGCGCTAGCCTGGCGCGAGTGCAGGTGCAGA 2530
 QY 1639 CGTCAAGGTACAGCTGCTGCTGAGCCCGAGGAGGCTCCAGTACCTGGTGGTATGCTTC 1698
 Db 2531 CGGTCAAGGTACAGCTGCTGCTGAGCCCGAGGAGGCTCCAGTACCTGGTGGTATGCTTC 2590
 QY 1699 GGCAGAGCCCGCGCTGACGACGCTGATCTAGCGGCTGCCAACTGCGCCGCCCATGG 1759
 Db 2591 GGCAGAGCCCGCGCTGACGACGCTGATCTAGCGGCTGCCAACTGCGCCGCCCATGG 2650
 QY 1759 TGACCTACCTGTGGAGTCTTCAGACACGAGGATGCGGCTTGCAGACCTCAGTCT 1816
 Db 2651 TGACCTACCTGTGGAGTCTTCAGACACGAGGATGCGGCTTGCAGACCTCAGTCT 2708

RESULT 8
 ABAC8979
 ID ABA08979 standard; cDNA; 1795 BP.
 AC ABA08979;
 XX
 D* 11-JAN-2002 (first entry)
 XX
 DE Human vasopressin receptor homologue-encoding cDNA, SEQ ID NO:755.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; anti-inflammatory;
 KW antiaesthetic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytotatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnary; antitumor; ss.
 OS Homo sap:ens.
 XX
 XX
 PN #0200157188-A2.
 XX
 XX 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX
 PI Tang YF, Liu C, Drmanac RT;
 XX
 XX WPI; 2001-457740/49.
 DR P-PSDB; ABB11735.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 XX Cialm 1; Page 678-679; 1963pp; English.
 XX
 CC Sequences ABB:0981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which

CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.
 XX
 SQ Sequence 1795 BP; 297 A; 567 C; 656 G; 275 T; 0 other;
 Query Match 50.9%; Score 1368.6; DB 22; Length 1795;
 Best Local Similarity 99.7%; Pred. No. 7.9e-236;
 Matches 1371; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 CGTGGCGCGGAGCTGCTGCTGACCTGATCCTGGACCACTGCCCGACCGCGCGCGCGG 66
 Db 421 CGGGCAGCGCAGCGCTGCTGCTGACCTGATCCTGGACCACTGCCCGACCGCGCGCGG 480
 QY 67 TGCCCGCAGATGCTGGCGCGCGCGCGCGCGCTCTTTCATCTCTGACGCGCGGACGAGC 126
 Db 481 TGCCCGCAGATGCTGGCGCGCGCGCGCGCGCTCTTTCATCTCTGACGCGCGGACGAGC 540
 QY 127 TGCCCGCGCTGGGGGCGCGCGCGCGCGCGCTCTTTCAGACAGACCCCTTTCGAGCGCGGAGCG 186
 Db 541 TGCCCGCGCTGGGGGCGCGCGCGCGCGCGCTCTTTCAGACAGACCCCTTTCGAGCGCGGAGCG 600
 QY 187 GCGCGCGGCTGCTAGCGCGGCTGCTGAGTAGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 246
 Db 601 GCGCGCGGCTGCTAGCGCGGCTGCTGAGTAGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 247 TGACCAACGCGCGCGCGCGCGCGCGCGCGCGCTGCTGAGTAGGCGGCTGCTGCTGCTGCTG 306
 Db 661 TGACCAACGCGCGCGCGCGCGCGCGCGCGCGCTGCTGAGTAGGCGGCTGCTGCTGCTGCTG 720
 QY 307 CCGAGGTGCGGGCTTCTCCGACAAAGGACAAAGAAAGTATTTCTCAAGTTCCTCCGGG 366
 Db 721 CCGAGGTGCGGGCTTCTCCGACAAAGGACAAAGAAAGTATTTCTCAAGTTCCTCCGGG 780
 QY 367 ATGAGAGAGGCGCGAGCGCGCTTACCGCTTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 426
 Db 781 ATGAGAGAGGCGCGAGCGCGCTTACCGCTTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 840
 QY 427 TGTGCTTGTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
 Db 841 TGTGCTTGTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 QY 487 TCGGTGCGGAGCTGCTGCGGCGAGCTCCAAAGAGACACACAGAGAGAGAGAGAGAGAGAGAG 546
 Db 901 TCGGTGCGGAGCTGCTGCGGCGAGCTCCAAAGAGACACACAGAGAGAGAGAGAGAGAGAGAG 960

Db 359 AAACGTGGTACACTACAGTTCATTCAGCAGAGCTTCCAGAGTCTTCTGCTGCATGT 418
Qy 805 CTTACCTGCTGGAGAGCGCGGGTGCCAGGACCGCGGCTGGCGGCTTGGGACACTCC 864
Db 419 CATACCTACTAGAGCTGAGAGGAGCCCGAGGAACTCCGAGGAAGTGTGAGATGCTCC 478
Qy 865 TGGCTGGGGAGCCCGAGCGGACAGCCACTTGGTGTCTCACACCGCTTCTCTTCCGAC 924
Db 479 TGCTCTCTGAGCGGGGCTGGTGTCTATCTGGCACTCACACCTAGATCTCTCTTTGGAC 538
Qy 925 TGCTGAGCGGGAGCGGATGCGGACATCGAGCGCACTTGGCTGTGATGTTTCAGAGC 984
Db 539 TGCTAAGTACAGAGAGGATTCGTGACATTGGAACCACTTGGCTGTGCTGGTCCAGGCG 598
Qy 985 GTGTGAAGCAGGAGCGCTCGGTTGGTGGTGAGGACAGAGGCG---TGCCCCGGAG 1041
Db 599 GTGTGAACAGGACACCTTGGGTTGGGTACAGGACAAAGCCAAAGGTGCGGACAG 658
Qy 1042 TGCCACACAGAGGTGACCGAGGGGGCCAAAGGGCTCGAGGACACCGAAGAGCGAGGAGG 1101
Db 659 TAGAGCAGAAAGAGAGGATGAGCTGAAGACGAGGAGAGAGAGGAGGAGGAGG 718
Qy 1102 AGGAGGAGGAGAGAGCCCAACTACCACTGGAGTTGCTGTACTGTGCTGTGAGAGCG 1161
Db 719 AAGAAGAGGAGGAGAACTCAACTTGGACTGGAGCTGTGTACTGTGCTGTATGAGACAC 778
Qy 1162 AGGAGGAGCGGTTGTGCGCCAAAGCCCTGTGCGGGTCCCGAGCTGCGCTGAGAG 1221
Db 779 AAGAGGATGATTTTGTTCGCGAGGCTCTCAGCAGCCCTCCAGAGATGATGATGAGGAG 839
Qy 1222 TGCGCTTCTGCGGATGAGCGTGGCTGTCTGAGCTACTGCGTGAGGTGCTGCGCTCTG 1281
Db 839 TTAGGCTACCGATGAGCTTGGAGTTCTGAGCTACTGCTGAGTGTGCTGCGGAGC 899
Qy 1282 GACAGGCACTCGCGTGTACAGTGTGAGTGTGCTGCGCAGGAGAGAGAGAGAG 1341
Db 899 GCCAGGCTCTGAGACTGTGTGAGCTGTGAGCTGGTGGCGGCAAGAGAGAGAGAG 958
Qy 1342 SCCTGGGAGAGGGCTCAGCGCAGCTGGGTGGGGCAGTCTCAAGGACACCAAAAC 1401
Db 959 -----AGAGAGCTTCAAGACCGGCTGAAG-----GTCTCAAGCACCAGGAAAC 1096
Qy 1402 AACTGCCAGCTCCCTCTGTATCACTCTTTTACGGAATGACTGACCACTGTGCCATC 1461
Db 1007 AACCCCCAGCTCTGTGCTGCTGCACCTGTGAGGCAATGATACCCAGCAATGTGCTC 1066
Qy 1462 TGAGCAGCTCACGCTGTCCCACTGCAAACTCCCTGACGCGGTCTGCGAGACCTTTCTG 1521
Db 1067 TGAGTATCTGACCTGTGCACACTGCAAACTCCCTGATGCAGTTTGTGAGACCTTTCTG 1126
Qy 1522 AGGCTCAGGGCAGCCCCGACCTGACGAGCTGGGGCTCTCTCAACAGGCTCAGTG 1581
Db 1127 AGGCTCTGAAGGTASCTCTTCCCTAAGGAGCTGGGGCTCTCTCAGAACCGACTCACTG 1186
Qy 1582 AGGCGGAGCTGATGCTGAGTAGGCGCTAGCTGGCGCCAGTGCAGGCTGAGAGCGG 1641
Db 1187 AGGCGGAGCTGGGTTTATTAAGCAGAGCGCTGGCTTGGCCCAATGCAAGGTGAGACAC 1246
Qy 1642 TCAGGATACAGTGTGCTGACCCCAAGCGAGGCTCCAGTACCTGGTGGGTATGCTTCGGC 1701
Db 1247 TCAGGATACAGTGTGCTGGGCTCCAAAGAGGTGATCCACTACTGTGTCTTGTCTCAGC 1306
Qy 1702 AGAGCCCCGCTGACCACTCTGGATCTCAGCGGCTGCCAACTGCCCGCCCCCATGGTGA 1761
Db 1307 AGAGCCCCAGTCTCAACCACTCTGCACTCAGTGGCTGTGAGTGTCTGGGAGTCTGGTGG 1366
Qy 1762 CTAAGCTGTGCTGCTGCTGAGCAGGAGGCTGCGCTGAGACCTCAGTCTGCTGCT 1821
Db 1367 AACCTCTGTGTTACAGCCCTCAACACCCCTAAATGTGCGCTTAAGACCTCTAGTCTGACTT 1426
Qy 1822 CTGTGAGCTGAGCGGAGCACTCACTACAGGAGCTTCAGGCTGTGAAGAGAGCAAGCGG 1881

Db 1427 CTGTGAGCTGACTGAGAAATCCACTGAGAGAGCTTCAAGCTGTGAAGACATTAAGCCAG 1486
Qy 1882 ATCTGCTCATCACACACCAGCGCTGGAGCGGCCACCAAACTCCCAAGGAATCATCT 1941
Db 1487 ATCTGGCCATCATACATTCAAAATTTGGGCACACATCTCAGCTCTGAAGGGATGAAGCA 1546
Qy 1942 CGACCTTCTTGAGGCTCTGGTGGCCAGAGCAGGCTGGAAGACCTAGTCAAAAGTCCCTGTG 2001
Db 1547 GTCTCTCTGA-GCCCCGAGTCCAGAACAGAT--ATTGCTGAAGTCAACGTTCTTCTG 1603
Qy 2002 GAGAGAACCGCCATTCAAAGGCGAGGAGGATATTGC-TCTCGGCTTTGGGAAACTTTT 2060
Db 1604 AATACTCTCTCCCATTCAAAAGCAAGTAGATGCCAGCTTCTCCCTCCAGGCAAGTCTTTC 1663
Qy 2061 GAGCCGAGAGGCGGACAGCAGGCAATGTGGAGGCCAG 2098
Db 1664 AAGCCAAAGGCCACAGAAGGGCAAGAAAAGACCCAG 1701

RESULT 10
AAK83313
ID AAK83313 standard; DNA; 933 BP.
XX AAK83313;
AC AC
XX XX
DT 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38125.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
XX
PK WC2C3157182-A2.
XX C9-AUG-2001.
PD
XX
PF 17-JAN-2001; 20C1WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 24-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216847.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0229924.
PR 01-SEP-2000; 2000US-0229257.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 31-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure; SEQ ID NO 38125; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 933 BP; 146 A; 309 C; 325 G; 153 T; 0 other;

Query Match 29.7%; Score 798.6; DB 22; Length 933;
Best Local Similarity 99.5%; Pred. No. 7.3e-134;
Matches 801; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 7 CGTGGCGCGCAGCCTGGCTGACCTGATCTGGACCAAGTSCCCGACCGCGCGCGCG 66
DB 129 CGGCGACGCGCAGCCTGGCTGACCTGATCTGGACCAAGTSCCCGACCGCGCGCGCG 188
QY 67 TGCCCGAGATGCTGGCGCAGCGCGCGCTGCTTCTTCATCTGTGACCGCGCGAGGAC 126

PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240969.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246471.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-483426/52.
DR P-PSDB; AAM89960.
DR
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -

XX Claim 1; SEQ ID NO 7801; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I);
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients' own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I) by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 989 BP; 157 A; 317 C; 344 G; 164 T; 7 other;
SQ
Query Match 27.5%; Score 740.6; DB 22; Length 989;
Best Local Similarity 94.5%; Pred. No. 1.8e-123;
Matches 795; Conservative 3; Mismatches 39; Indels 4; Gaps 3;
QY 7 CGTGGCGCGCAGCTGGCTGACCTGATCTCTGGACAGCGCCGACCGCGCGCGCGCGG 66
DB 129 CGGCGACGCGCAGCTGGCTGACCTGATCTCTGGACAGCGCCGACCGCGCGCGCGG 198
QY 67 TGCCCGCAGATGCTGGCCCGCAGCGCGCTCTTTCATCTCTGACGCGCGCAGCAGC 126
DB 189 TGCCCGCAGATGCTGGCCCGCAGCGCGCTCTTTCATCTCTGACGCGCGCAGCAGC 248
QY 127 TSCCGGCGCTGGGGCGCGCGAGCGCGCCCTGTCACAGACCCCTTCGAGGGCGGAGCG 186
DB 249 TGCCCGGCGCTGGGGCGCGCGAGCGCGCCCTGTCACAGACCCCTTCGAGGGCGGAGCG 308
QY 197 GC CGCGCGGGTGCTAGCGGGGCTGCTGAGTAAGCGCTGCTGCCACGCGCCCTCTCTGCTGG 246
DB 309 GC CGCGCGGGTGCTAGCGGGGCTGCTGAGTAAGCGCTGCTGCCACGCGCCCTCTCTGCTGG 368
QY 247 TGACCAAGCG 306
DB 369 TGACCAAGCG 422
QY 307 CC GAGGTGGCGGCTTCTCCGACAAAGGACAAAGAAAGTATTTCACAAGTTCTTCCCGG 366
DB 429 CC GAGGTGGCGGCTTCTCCGACAAAGGACAAAGAAAGTATTTCACAAGTATTTCGCGG 488
QY 367 ATGAGAGGAGGCG 426
DB 489 ATGAGAGGAGGCG 548
QY 427 TGTGCTTCGTCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
DB 549 TGTGCTTCGTCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 608
QY 497 TC GGTTCGGACCTGTTCGCGCAGCTCAAGACCAACACCTCAGTGTACCTGCTTTCATCA 546
DB 609 TCGGTTCGGACCTGTTCGCGCAGCTCAAGACCAACACCTCAGTGTACCTGCTTTCATCA 668
QY 547 CCAGCGCTTCGAGCTCGGCTCGGCTAGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
DB 669 CCAGCGCTTCGAGCTCGGCTCGGCTAGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCG 728
QY 607 ATCTGTCCCGCTTGCG 666
DB 729 ATCTGTCCCGCTTGCG 788
QY 667 AACTGAGCAACTGAGCTTCGTGCTCG- AAGTGTGACAGCG- TGTTTCTCAGCAAAA 724

Db 789 AACTGGAGCAACTGGAGCTTCGTGGCTCCAAAAGTGCAGAGCTGTGTTCTCAAGAAAA 848

Qy 725 GGAGTCGCCGGCGTGGTGGAGACAGAGTCACTACCAAGTTCACTCGACCA--GAGCTTC 782

Db 849 GGAAGTGGCGGGCGTGGTGGANACAGAGTCACTACCAAGTTCACTCGACCAAGAGCTTTN 908

Qy 783 CAGGAGTTCTTCGGGGCACTGCTCTACCTGTGGAGGACGGCGGGTGCCCGAGGACCGCG 842

Db 939 CAGGAGTTCTTCGGGGCACTGCTCTACCTGTGGAGGACGGCGGGTGCCCGCAAG 968

Qy 843 G 843

Db 969 G 969

RESULT 12

ABL90582

ID ABL90582 standard; cDNA; 990 BP.

XX

AC ABL90582:

XX

CT 24-MAY-2002 (first entry)

XX

XX Human polynucleotide SEQ ID NO 1144.

DE

XX Cy-tostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;

KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; human; secreted protein; gene; ss.

XX

OS Homo sapiens.

PN WO200190304-A2.

XX

XX 29-NOV-2001.

PD

XX 18-MAY-2001; 2001WO-US16450.

PF

XX 19-MAY-2000; 2000US-205515P.

PR

XX (HUMA-) HUMAN GENOME SCI INC.

PA

XX Birse CE, Rosen CA;

PI

XX WPI: 2002-122019/16.

DR

XX P-PSDB; ABB90173.

XX

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and

PT prevention of neural, immune system, muscular, reproductive,

PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders -

XX

XX Claim 4: SEQ ID NO 1144; 2081pp + Sequence Listing; English.

PS

XX The invention relates to novel genes (AB189449-AB190853) and proteins

CC (AB189040-AB190444) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are

CC isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful

CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast

CC and ovarian cancer and other cancers of the adrenal gland, bone, bone

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's

CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;

CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and

CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal

CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 990 BP; 158 A; 318 C; 344 G; 163 T; 7 other;

Query Match 27.1%; Score 729.6; DB 24; Length 990;

Best Local Similarity 94.4%; Pred. No. 1.6e-121;

Xatches 795; Conservative 3; Mismatches 39; Indels 5; Gaps 4;

Qy 7 CGTGGCGCGCAGCCTGGCTGACCTCATCTGTGACCAAGTCCCGGACCGCGGCGCGCG 66

Db 129 CGGGCAGCGCAGCCTGGCTGACCTGATCTGTGACCAAGTCCCGGACCGCGGCGCGCG 188

Qy 67 TCGCGCAGATGTGTGCGCCAGCCGACGCGCTCTTTCATCTCTGACGCGGGACGAC 126

Db 189 TCGCGCAGATGTGTGCGCCAGCCGACGCGCTCTTTCATCTCTGACGCGGGACGAC 248

Qy 127 TCGCGCGCTGGGGGCG 186

Db 249 TCGCGCGCTGGGGGCG 308

Qy 187 GCGCGCGGCTGTAGCGCGCTGTGCTAGTAAGCGCGCTGTGCGCCACGCGCGCTCTCTG 246

Db 309 GCGCGCGGCTGTAGCGCGCTGTGCTAGTAAGCGCGCTGTGCGCCACGCGCGCTCTCTG 368

Qy 247 TGACACG 306

Db 369 TGACACG 428

Qy 307 CGAGGTGCGCGCTCTCGGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 366

Db 429 CCAGGTGCGCGCTCTCGGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 428

Qy 367 ATGAGAGGAGCGCGCGCGCGCGCGCGCGCTACCGCTTCGTGAGGAGCAAGCAAGCAAG 426

Db 489 ATGAGAGGAGCGCGCGCGCGCGCGCGCTACCGCTTCGTGAGGAGCAAGCAAGCAAG 548

Qy 427 TGTGCTTCGTGCGCTTCGTGCTGGATCGTGTGACCGCTGTGCGCGCGCGCGCGCGCG 486

Db 549 TGTGCTTCGTGCGCTTCGTGCTGGATCGTGTGACCGCTGTGCGCGCGCGCGCGCGCG 608

Qy 487 TCGGTGCGGACCTGTGCG 546

Db 609 TCGGTGCGGACCTGTGCG 608

Qy 547 CCAGCGCTTCGTGAGCTCGGCTTCGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606

Db 669 CCAGCGCTTCGTGAGCTCGGCTTCGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 728

Qy 607 ATCTGTCGCGCTTCGTGCG 666

Db 729 ATCTGTCGCGCTTCGTGCG 788

Qy 667 AACTGAGCAACTGGAGCTTCGTGCGCTCC-AAAGTGCAGCGC-TGTTTCTCAGCAAAA 724

Db 789 AACTGAGCAACTGGAGCTTCGTGCGCTCCAAAGTGCAGCGCTTCCTTCTCAGCAAAA 848

Qy 725 GGAGCTG-CGCGCGCTGGAGACAGAGTCACTACCAAGTTCACTCGACCA--GAGCTTC 781

Db 849 GGAGCTGCGCGCGCTGGAGACAGAGTCACTACCAAGTTCACTCGACCAAGAGCTTC 908

Qy 782 CAGGAGTTCTTCGGCGCACTGTCTTCTGCTGTGAGGACCGCGGGGTGCCCGAGCGCG 841

Db 909 NCAGGAGTTCTTCGGCGCACTGTCTTCTGCTGTGAGGACCGCGGGGTGCCCGAGCG 968

Qy 842 GG 843

Db 969 GG 970

RESULT 13

ABQ29970

ID ABQ29970 standard; DNA; 704 BP.

XX

AC ABQ29970;

```

XX CC
XX CC
XX DE 12-JUL-2002 (first entry)
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 16562.
XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX OS Homo sapiens.
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EPI0074.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PR 05-SEP-2000; 2000DE-1044543.
XX PA (EPIG-) EPIGENOMICS AG.
XX FI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX DR WPI; 2002-371829/40.
XX PT Determining the degree of cytosine methylation in genomic DNA, useful
XX PT for diagnosis and prognosis, comprises selective hybridization of
XX PT amplicons from chemically treated DNA
XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX CC
XX CC This invention describes a novel method for determining the degree of
XX CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
XX CC genomic sample of DNA. The sample is treated chemically to convert
XX CC cytosine (C) but not methylated C to uracil, then part of the genomic
XX CC DNA that contains the target C is amplified to form a labeled amplicon.
XX CC The amplicon is hybridised to two classes, each with at least one
XX CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
XX CC and the degree of hybridisation to both classes is determined from the
XX CC label on the amplicon. From the ratio of labels hybridised to the two
XX CC classes of oligomers, the degree of methylation is calculated. The method
XX CC is used: (i) for diagnosis and/or prognosis of side effects of
XX CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
XX CC of the central nervous, cardiovascular, gastrointestinal, and respiratory
XX CC systems etc., particularly by detecting mutations or single nucleotide
XX CC polymorphisms (SNPs); and (ii) for differentiation of cell or tissue
XX CC types and for investigating cell differentiation. The method allows the
XX CC methylation status of many C residues to be determined simultaneously.
XX CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
XX CC method for determining the degree of cytosine methylation described in
XX CC the disclosure of the invention.
XX SQ Sequence 704 BP; 106 A; 82 C; 251 G; 265 T; 0 other;
XX CC
XX CC Query Match 17.3%; Score 464.6; DB 24; Length 704;
XX CC Best Local Similarity 78.8%; Pred. No. 4.1e-74;
XX CC Matches 554; Conservative 0; Mismatches 49; Indels 2; Gaps 0;
XX QY 50 CGACCGGGCGCGCGGTGCGGAGATGCTGGCCACCGCGCGCTGCTTCATCT 109
XX DB 2 CGATCGGGCGCGTGGGTGCTAGATGTTGGTTAGTGGGTGGTTTATTTT 61
XX QY 110 GGACGGCGGACGAGCTGCGGGCTGGGGGCGCGCGCGCGCGCGCGCGCACACCC 169
XX DB 62 GGACGGCGGACGAGTGTGCGGGTGGGGGGTTTCGAGGTCCGCTTTGTATAGATT 121
XX QY 170 CTTGAGCGCGGCGCGCGCGCGGTGCTAGCGGGCTCTGAGTAAGCGCTGCTGCC 229
XX DB 122 TTTGAGGCGCGGAGCGCGCGGGTGTAGCGGGTGTGAGTAAGCGGTGTGT 181
XX QY 230 CACGGCGCTCTCTGTGTGACCAACGCGCGCGCGCGCGCGCGCGCGCGCGCT 289

```

```

DB 182 TACGGTTTTTTTGTGTTGATTACGGCGCTGCTGTTTTCCGGAGGTTGTCAGGTCGTTT 241
QY 290 GTGTTCCCGCAGTGGCGGAGTGCGCGGCTTCTCCGACAAAGGACAGAAAGTATT 349
DB 242 GTGTTTTTGTAGTGGTGGCGGTGTTTTTTCGATAAGGATAAGAAAGTATT 301
QY 350 CTACAAAGTCTTCCGGGATGAGAGGAGGCGCGGCTACCGTTCGTCAAGGAGAA 409
DB 322 TTATAAGTTTTTCCGGATGAGAGGAGGCGGTTCGTTATCGTTTCGTGAAGAGAA 362
QY 410 CGAGACGCTGTTCCGCGTGTGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 469
DB 362 CGAGACGCTGTTCCGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
QY 470 GCGCCAGCAGCTGAGCTGCGGACCTGTCGCGCAGTCCGAGACACACACGTCAGT 529
DB 422 GCGTTAGTAGTTGGAGTTCGGTCGGGACTTGTGCGGTAGCTTTAAGATTATTACGTTAGT 481
QY 530 GTACCTGCTTTTTCATCACCAGCGTCTGAGCTCGGCTCCGCTAGCGCGCGCGCGGTT 589
DB 482 GTATTTGTTTTTATTATTATTAGCGTTTTCGTTTCGGTAGTCGACGGGTTCGGTT 541
QY 590 GCAGGGCGACCTCGGCAATCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
DB 542 GTAGGGCGGATTTTCGTAATTTGCTGCTTGGTTCCGAGGGCGGTTTTCCGAGCGTAGGGC 601
QY 650 GCAGTTTCCGAGAGGAACTGGAGCACTGGAGCTTCGTTGGCTCCAAAGTCGACAGCT 709
DB 602 GTAGTTTGTGAGAGGAACTGGAGTAATTTGGAGTTTCGTTGGTTTAAAGTGTAGACGTT 661
QY 710 GTTTCTCAGCAAAAAGGAGCTGCGGGCGCTGCTGGAGACAGAG 752
DB 662 GTTTTTTAGTAAAAAAGGAGTTGTCGGGCGTGTTCGAGATAGAG 704
XX RESULT 14
XX ABQ29971/C
XX ID ABQ29971 standard; DNA; 704 BP.
XX AC ABQ29971;
XX XX
XX DT 12-JUL-2002 (first entry)
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 16562.
XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX OS Homo sapiens.
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EPI0074.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PR 05-SEP-2000; 2000DE-1044543.
XX PA (EPIG-) EPIGENOMICS AG.
XX FI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX DR WPI; 2002-371829/40.
XX PT Determining the degree of cytosine methylation in genomic DNA, useful
XX PT for diagnosis and prognosis, comprises selective hybridization of
XX PT amplicons from chemically treated DNA
XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

```



```
QY 109 TGAACGGCGCGACGAGCTGCGCGGCTGGGGGGCCCCAGAGCCGCGCCCTGCACAGACC 168
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
644 TAAACGACGCGAACGAACTACCGACGCTAAATAACCCCGAAACCCGCGCCCTACACAAACC 585
QY 169 CCTTCGAGGCGCGAGCGCGCGGTGCTAGGCGGGCTGCTAGAGCGCTGCTGC 228
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
584 CCTTCGAAACGACGACGCGCGAATACTAAACGAATCTTAAACAAACGCTACTAC 525
QY 229 CCACGGCCCTCTGCTGCTGCTCACCACGCGCGCCCGCGGAGGCTGCAGGGCCGCC 286
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
524 CCACGACCTCTCTACTAAATACCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 465
QY 289 TGTGTTCCCGGAGTGGCGCGAGGTGGCGGCTTCTCCGACAGGACGAGAGAGATAT 348
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
464 TATATTCGCCGCAATACCGCGAATACGCGACTTCTCCGACAAACCAAAAAAATAAT 405
QY 349 TCTACAAATTTCTTCGGAATAGAGGAGGCGCGAGCGCTTACCGCTTCGTGAAGGAGA 408
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
404 TCTACAAATTTCTTCGGAATAGAGGAGGCGCGAGCGCTTACCGCTTCGTGAAGGAGA 345
QY 409 ACGAGAGCTGTTCGGGCTGTGCTTCGTGCCCTTCGTGTGATCGTGTGACCGTGC 468
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
344 ACGAAACGCTATTTCGGCTATCTTCGTACCTTCGTATATAAATCGTATACCGGTAC 285
QY 469 TGGCGGAGGAGCTGGAGCTCGGTGCGGACCTGTCCGCGACGTCCAGACGACGAGTCAG 528
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
284 TACGCCAACAACTAAACCTCGATCGAAACCTATCGCGCACGTCCCAAAACCAACGTCAA 225
QY 529 TGTACCTGCTTTTTCATCACCAGCGTTCTGAGCTCGGCTCGGTAGCGGACGGCGCCCGT 586
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
224 TATACCTACTTTTTCATCACCAGCGTTCTGAACTCGACTCGATACCGAGCAACCCGAT 165
QY 589 TGCAGGCGGACCTGCGCAATCTGTGCGGCTGCGCGCGAGGCGCTCTCGGACGCGAGG 648
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
164 TACAAAACGACCTACGCAATCTATACGCGCTAACCCGCGGAAACGCTCTCGAACGCAAAA 105
QY 649 CGCAGTTTGCAGAGAGGAACTGGAGCACTGGAGCTTCGTGGCTCCAAAGTGCAGACGC 708
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
104 CGCAATTTACCGAAATAAATACTAAACAACTAAATACTTGTAACTCCAAATAACAAACGC 45
QY 709 TGTTCACGAAAGAGAGCTCGCGGCGTGTGGAGACAGA 751
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
44 TATTTCTCAACAAAAAAAATCTACCGAACGTAATAAAACAAA 2
```

Search completed: October 3, 2003, 06:57:34
Job time : 703 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - nuc-eic search, using frame_plus_p2n.model

Run on: October 3, 2003, 13:31:41 ; Search time 3480 Seconds
(without alignments)
4365.028 Million: cell updates/sec

Title: US-10-029-347-2

Perfect score: 3218

Sequence: 1 MLAQPCRLFLIDGABELPA.....ITHPALDGHPOPFKELISTF 625

Scoring table:

BL-OSUM62
Xgapcp 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12:52238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEJ=frame+p2n.model -DEV=xlh
-Q=/cgn2/USPTO_spool/US10029347/runat_02102003_173436_17085/app_query.fasta_1.775
-DE=EST -Qfmt=fastap -SUFFIX=est -MINMATCH=0.1 -LOPCU=0 -LOCPXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pcp -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029347 -CGN 1 1 2810 -runat_02102003_173436_17085 -NCFU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -JONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=0 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:
2: em_esthum:
3: em_estin:
4: em_estmu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_hci:
9: gb_est1:
10: gb_est2:
11: gb_hci:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estom:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pln:
20: em_gss_vrt:
21: em_gss_fun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_rod:
26: em_gss_phg:
27: em_gss_vrl:
28: gb_gss1:

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2124	66.0	2730	11	AK039531 Mus muscu
2	927.5	28.8	798	9	AW107005 um18d07.y
3	901	28.0	719	12	BG971110 602837817
4	830	25.8	691	9	AI344276 tc03a07.x
5	808	25.1	938	10	BF783992 602107835
6	793	24.6	603	12	BG971489 602838184
7	73.5	22.7	832	12	BG965959 602829753
8	603.5	18.8	2218	11	AF054176 Homo sapi
9	601.5	18.7	416	13	BY262618 BY262618
10	601	18.7	670	10	BB633515 BB633515
11	598.5	18.6	3368	11	BC021272 Homo sapi
12	598.5	18.6	3369	11	BC012789 Homo sapi
13	499	15.5	3400	11	AK054264 Mus muscu
14	487	15.1	431	13	BY281505 BY281505
15	473	14.7	3577	11	AK087843 Mus muscu
16	466	14.5	3218	11	AK054378 Mus muscu
17	458.5	14.2	3475	11	AK087774 Mus muscu
18	444.5	13.8	3793	11	BC036506 Homo sapi
19	433	13.5	394	10	BF561521 UI-R-CO-H
20	412.5	12.8	3275	11	AK054426 Mus muscu
21	406	12.6	488	9	AI037137 uc90h05.y
22	390	12.1	791	28	B2129303 CH230-345
23	381.5	11.9	1110	13	BA57482 BX457482
24	362.5	11.3	377	9	A1558213 vj37a07.y
25	357.5	11.1	939	13	BA371966 BX371966
26	356	11.1	2520	11	AK032446 Mus muscu
27	356	11.1	4476	11	AK028802 Mus muscu
28	354	11.0	2394	11	AK028575 Mus muscu
29	352.5	11.0	664	29	AG067278 Pan trogl
30	339.5	10.6	921	13	BQ431649 AGENCOURT
31	318	9.9	532	9	AA541951 vj37a07.r
32	315	9.8	4442	11	AK089843 Mus muscu
33	314	9.8	595	9	A1883819 fc70d12.y
34	313	9.7	710	28	BH065379 RPI-24-2
35	310.5	9.6	3771	11	BC023974 Homo sapi
36	304.5	9.5	577	28	AZ432279 IM22170C8
37	297.5	9.2	791	10	BG682080 602630057
38	286	8.9	934	10	BE733220 601568455
39	283	8.8	629	14	CB016418 pgnlc.pk0
40	281	8.7	1156	10	BE731237 601567129
41	278	8.6	2603	11	AK089390 Mus muscu
42	276	8.6	588	14	CA559792 K0263E12-
43	271.5	8.4	1114	14	CD506212 CDA78-A06
44	266.5	8.3	646	10	BE265950 601193029
45	262	8.1	639	28	AZ587775 IM0395309

ALIGNMENTS

RESULT 1
AK039531
LOCUS 2730 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus adult male spinal cord cDNA, RIKEN full-length
enriched library, clone:A330055K17 product:similar to VASOPRESSIN
RECEPTOR [Rattus norvegicus], full insert sequence.
ACCESSION AK039531
VERSION AK039531.1 GI:26087232
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636

REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11942159

REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Hatahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861

REFERENCE

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, C., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadoya, K., Matsuda, H., Ashburner, K., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schrim, L. M., Stauble, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marzocchi, L., Mashima, C., Mazzatelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kotsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection:
Nature 409 (6821), 685-690 (2001)
MEDLINE
21085660
PUBMED
11927831

REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase 1 & 11 Team.
Analysis of the mouse transcriptome based on functional annotation:
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
MEDLINE
11942159
PUBMED
11942159

REFERENCE

6 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuo, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, Y.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Yuramatsu, M. and Hayashizaki, Y.
Direct Submission

REFERENCE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome

REFERENCE

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

Location/Qualifiers
source
1..2730
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="TANTOM:DB:A330055X17"
/db_xref="taxon:10090"
/clone="A330055X17"
/sex="male"
/tissue_type="spinal cord"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1..2730
/note="similar to VASOPRESSIN RECEPTOR [Rattus norvegicus]
(SPTRIQ63035, evidence: FASTY, 82.2%ID, 100%length,
match=1440)"

BASE COUNT 649 a 730 c 767 g 584 t

ORIGIN

Alignment Scores:
Pred. No.: 3,328-193 Length: 2730
Score: 2124.00 Matches: 433
Percent Similarity: 79.71% Conservative: 66
Best Local Similarity: 69.17% Mismatches: 109
Query Match: 66.00% Indels: 18
DB: 11 Gaps: 8

US-10-029-347-2 (1-625) x AK039931 (1-2730)

Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuLeuAspGlyValaAspGluLeuProAla 20
:::|||||

Db 233 ATACTGGCGCACCCCAATCGCTACTGTTCATCTGGAGCGTGCAGACGAGTGGCTACT 292
|||

Qy 21 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyalaArg 43
|||

Db 293 TTACCATCTCTCAGAGGCCACTCCCTGCCAAGACCCCTTTGGAGGCTACAAGTGGCTTGAGA 352
|||

Qy 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
|||

Db 353 GTGCTGAGCGGCTGCTGAGCAGGAGGCTGTTCACGAGGAGTGGCTTGTAGTACTAC 412
|||

Qy 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
|||

Db 413 CGCCATGCGCTACAGGAGGCTGCGAGCGCAGTGTGCTGCCACAGTGCAGAGATA 472
|||

Qy 81 ArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLysLys 100
|||

Db 473 CGCGGCTTCTCAGACAAAGACAAACAAAAAGTATTCTTCAAGTTCTTCGGGACGAGG 532
|||

Qy 101 ArgAlaGluArgAlaTyArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
|||

Db 533 AGGCGAGCGCGCTTACCGCTTCGTGAAGACAGACGAGCGCTGTTCGACTGTGCTTT 592
|||

Qy 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
|||

Db 593 GTGCCCTCTGTGCTGGATCGTGTGCAGTGTGCAGTGTGCAGCAGCAGTGCAGTGGCGG 652
|||

Qy 141 AspLeuSerArgThrSerLysThrThrThrSerValTyLeuLeuPheLeuThrSerVal 160
|||

Db 653 GGTCTCTCTCGTACCTCCCAAGACCACTACTTCTGTGTACCTGCTTCTATCACCAGCATG 712
|||

QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 713 CTGAAGTCTCGAGC---ACCAAATGGACCCCGGTTCAGGAGAGCTTGAGAACGCCTGTGT 769

QY 181 ArgLeuAlaArgGluGlyVa-LeuGly---ArgArgAlaGlnPheAlaGluLysGluLeu 199
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 770 CGCTGCCCAAGGAGGCATCTCTGGATCATATAAGCACAAATTCTCAGAAGAGGACCTG 829

QY 200 GluGlnLeuGluAryGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeu 219
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 830 GAGAAATTGAAGCTTCGGGG-TCCAAGTTCCAGACAATAITTTCAAACAGAAGAAGATA 889

QY 220 ProGlyValLeuGluThrGluValThrTy-GlnPheIleAspGlnSerPheGlnGlutPhe 239
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 890 CCAGGAGTGTAAAANACTGAGTCCACTACAGTTCATTACCCAGAGCTTCCCAGGAGTTC 949

QY 240 LeuAlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGly 259
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 950 TTGCTGCATTGTTCATACCTACTAGAGCTGAGCGAACTCCGGGAGCCCGCAGGCGGT 1009

QY 260 ValGlyThrLeuLeuAryGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArg 279
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1010 GTSCAGAAGCTCCTGAACCTTGAGCGGAGCTACGTGTCATCTTGCACTGACACCCGA 1069

QY 280 PheLeuPheGlyLeuLeuSerAlaGluAryMetArgAspIlegLuArgHisPheGlyCys 299
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1070 TTCCTCTTTGGCCTGCTAAATATACAGAGGGCGTCTGTGACATTCGAAACCACTTTGSGCTG 1129

QY 300 MetValSerGluAryValLysGlnGuLaLeuAryTrpValGlnGlyGlnGly- 317
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1130 GTGGTCCGACATATGTAALAAAGACACCCTCGGGGGGTACAAGSACAGAGCCACCCC 1189

QY 318 GlnGlyCysProGlyValAla-ProGluValThrGluGyAlaLysGlyLeuGyAspThr 337
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1190 AAGGGGCCACGATAGGGGGCLAAAAAGACTGCTGAG-----CTGGAGGACATC 1237

QY 338 GlnGluProGluGluGluGluGluGlyGluGluPro---AsnTyrProLeuGylLeuLeu 356
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1238 GAGGACCAGAGGAGGAGGAGAGAGAGAGAGAGGAGCCTCAATTGTGAGCTGGAGCTGCTG 1297

QY 357 TyrCysLeuTyrGluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPhePro 376
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1298 TACTGCCTGTATGAGACACAGGAGGAGATTTGTTCACAGAGCTCTTCACAGCCTTCCA 1359

QY 377 GluLeuAlaLeuGlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCys 396
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1358 GAGATAGTACTGGAGCGAGTTCAGTTGACCGCATGAGACCTTGAACTTCGAATTACTGC 1417

QY 397 ValArgCysCysProAlaGlyGlnAlaLeuArgLeuIleserCysArgLeuValAlaAla 416
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1418 GTGCGAGTGTGCCAGATGCGCAGGCTCTGAGACTGCTGAGCTGTGGCTGTGGCGGCA 1477

QY 417 GlnGlu-----LysLysLysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGly 434
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1478 AAGGAAGAAGAAGAAGAAGAAGAGCCTGTTGAAGCGCTGGAAG----- 1522

QY 435 GlySerSerGlnGlyThrThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGln 454
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1523 ---GGCTCTCANAGCACCAACAAACACCCCGAGTCTCTTGTGGTGGTCCACTCTCTGAG 1579

QY 455 AlaMetThrAspProLeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuPro 474
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1580 ACAATGACTACCCCGAAATGTCATCTCAGTGTTCTGTATCTTTGTCCACTGCAGACTCCCT 539

QY 475 AspAlaValCysArgAspLeuSerGluAlaLeuAryAlaAlaProAlaLeuThrGluLeu 494
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1640 GATGCAGTTTTGCCGAGACCTTCCGAGGCCCTGAAAGGTAGCTCTGCGCCTTAAGGGAGTTG 1699

QY 495 GlyLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAla 514
 ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1700 GGCTCTCTCAGAGCCCGGCTCACCAACACAGGCCCTGCGTTTACTGTGCGAAGCCCTGGCT 1759

Qy	515	TpProGlnCysArgValGlnThrValArgValGlnLeuProAspProGlnArgGlyLeu	534
Db	1760	TGGCCCAAGTGCAGGTGAAGACACTCAGATGCAGCTGCCTGACCTCCCAAGAGGTGATC	1819
Qy	535	GlnTyrLeuValGlyMetLeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGly	554
Db	1820	AATTACTTGTGTCATCGTCTCCAGCAGAGCCCACTCTGACCACTCTGGACCTCAGTGGC	1879
Qy	555	CysGlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCys	574
Db	1880	TGTCAGCTGCTGGCGTCATGTGGAACTCTGTGTGGACCCCTGAAGCACCCCTAAATGC	1939
Qy	575	GlyLeuGlnThrLeuSerLeuAlaSerValGlnLeuSerGlnGlnSerLeuGlnGluLeu	594
Db	1940	TCCTTAAAGACCCCTCAGCTCTGACCTCTGTGGAGCTGAGTGAGAACTCACTGAGGGATCTT	1999
Qy	595	GlnAlaValIlyArgAlaLysProAspLeuValIleThrHisProAlaLeuAspGlyHis	614
Db	2000	CAAGCTGTGAAGACATCAAGCCAGATCTGTCCATCATATAT---TCAAAATGAAGCACA	2056
Qy	615	ProGlnProProLysGlu	620
Db	2057	CCTCACCGGCTCGGAAG	2074
RESULT 2			
LOCUS	AW107005		
DEFINITION	IM8407.Y1 Sugano mouse kidney mKia Mus musculus cDNA clone		
	IM8407:2192653 5' similar to TR:Q63035 Q63035 VASOPRESSIN RECEPTOR.		
	1. mRNA sequence.		
ACCESSION	AW107005		
VERSION	AW107005.1	GI:6077873	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
	Other ESTs: un18d07.xl		
	Contact: Marra M/WashU-NCI Mouse EST Project 1999		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63106, USA		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: mouseest@watson.wustl.edu		
	This clone is available royalty-free through LNL; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	MG1:1005105		
	Seq primer: custom primer used		
	High quality sequence stop: 503.		
FEATURES			
SOURCE			
	1. .798		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:2192653"		
	/sex="female"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/clone_lib="Sugano mouse kidney mKia"		
	/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII		
	(CACTGTG); Site 2: DraIII (CACTATGTG); 1st strand cDNA		
	was primed with an oligo(dT) primer		
	[ATGTGGCTTTTTTTTTTTTTT]; double-stranded cDNA was		
	ligated to a DraIII adaptor [TGTTCCTACTGG], digested		

QY 113 uThrLeuPheAlaLeuCysPheValProPheValCysTrpIleValCysThrValLeuAr 133
 Db 256 AACCTGTTTGCATGCTGCTTGTGGCCCTTCGTGCTGGATCGTGGACAGTCTGCTGCC 315
 QY 133 gGlnGlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTy 153
 Db 316 GCACGAGCTAGAGCTGGCCGGATCTCTCTCGTACCTCCAGACCACTACTTCTGTGTA 375
 QY 153 rLeuLeuPheIleThrSerValLeuSerSerAlaProValAlaAspGlyProArgLeuG 173
 Db 376 CCTCTCTTTCATCACCAGCATGCTGAAGTCTGCAGGC--ACCAATGGACCCCGGTTC 432
 QY 173 nGlyAspLeuArgAsnLeuCysArgLeuAlaArgGluGlyValLeuGly--ArgArgA 192
 Db 433 GGGAGACTGAGAACGGCTGTCTCCCTTGGCCAGGAGGGCATCTCGATCATCATTAAGC 492
 QY 192 aGlnPheAlaGluLysGluLeuGluGlnLeuGluLeuArgLysSerLysValGlnThrLe 212
 Db 493 ACAATCTCAGAGAGGACCTGGAGAAATCGAAGCTTCGGGGTTCCCAAGTTCAGACAA 552
 QY 212 uPheLeuSerLysLysGluLeuProGlyValLeuGluThrGluValThrTyrgInPhe 232
 Db 553 ATTCTCAACAGAGAGATACACAGAGTCTTAANAACCTGAGGTCACTTACCAGTTTCA 612
 QY 232 eAspGln-SerPheGlnGluPheLeuAlaLeuSerTyThrLeuLeuGluAspGlyGlyV 252
 Db 613 TGACCAAGAGCTTCAGGAGTCTTGCTGATGTCTATACCTACTAGAGAGCTGAGCGAA 672
 QY 252 alp-oArgThrAlaAlaGlyGlyValGlyThrLeuLeuA:GGiYasp 267
 Db 673 CTCGGGAGCGCCCGCAGGCGTGTGCGAAGAGCTCTGAACTCTGAG 719
 RESULT 4
 A1344276 691 bp mRNA linear EST 08-APR-1999
 LOCUS tc03a07.x1 NCI CGAP Col6 Homo sapiens cDNA clone IMAGE:206274C 3
 DEFINITION similar to TR:Q63035 Q63035 VASOPRESSIN RECEPTOR. ; mRNA sequence.
 ACCESSION A1344276
 VERSION A1344276.1 GI:4081482
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 691)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapb-r@mail.nih.gov
 Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.,
 Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnlnl.gov/bbrp/image/image.html
 Insert length: 1242 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 459.
 Location/Qualifiers
 1..691
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2062740"
 /tissue_type="colon tumor, RER-"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Col6"
 /note="Organ: Colon; Vector: p7T3D-Pac (Pharmacia) with a

modified polylinker; Site 1: Not I; Site 2: Eco RI;
 plasmid DNA from the normalized library NCI CGAP Col6 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (clones 1057416-1061255, and 1144584-1145351).
 Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 100 a 232 c 245 g 113 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1,936-69 Length: 691
 Score: 830.00 Matches: 162
 Percent Similarity: 99.39% Conservative: 0
 Best Local Similarity: 99.39% Mismatches: 1
 Query Match: 25.79% Indels: 0
 DB: 9 Gaps: 0
 US-10-029-347-2 (1-625) x A1344276 (1-691)

QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProhla 20
 Db 197 ATGCTGGCCCAAGCCGACGGCTGCTTTCATCTCTGGACGCGCGGAGAGCTGCCGCGC 256
 QY 21 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
 Db 257 CTGGGGGGCCCGAGGCGCGCCCTTCACACACCCCTTCGAGGCGCGCGCGCGCG 316
 QY 41 ValLeuGlyGlyLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
 Db 317 GTGCTAGGCGGGCTCTGAGTAAGCGCTGTCTGCCACCGGCCCTCTCTGTGTGACAC 376
 QY 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
 Db 377 CCG 435
 QY 81 ArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLysLys 100
 Db 437 CCGCGGCTTCTCCGACAGGACAAAGAAGATATTTCTACAGTTCTTCCGGGATGAGAG 496
 QY 101 ArgAlaGluArgAlaTyArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
 Db 497 ASGGCGGAGCGCGCGCTACCGCTTCGTGAAGGAGAACGAGACGCTGTCCGCGCTGTCT 556
 QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
 Db 557 GTGCCCTTCGT 616
 QY 141 AspLeuSerArgThrSerLysThrThrThrThrThrThrThrThrThrThrThrThrThr 160
 Db 617 GACCTGTCCGCGACCTCCCAAGACCAACGACGTCAGTGTACTGTCTTTCATCACCAG 676
 QY 161 LeuSerSer 163
 Db 677 CTGAGCTCG 685

RESULT 5
 BF783992
 LOCUS 602107835F1 NCI CGAP Kid14 Mus musculus cDNA clone IMAGE:4236272
 DEFINITION 5', mRNA sequence.
 ACCESSION BF783992
 VERSION BF783992.1 GI:12089028
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 938)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished

Score: 793.00 Matches: 160
 Percent Similarity: 90.10% Conservativity: 22
 Best Local Similarity: 79.21% Mismatches: 16
 Query Match: 24.64% Indels: 4
 DB: 12 Gaps: 2

US-10-029-347-2 (1-625) x BG971489 (1-603)

```

QY 31 AspPropheGluAlaA1a-SerGlyAlaArgValLeuGlyGlyLeuLeuSerIysAlaLe 50
DQ 3 GACCTTTGGAGGCTACAAATGTGCTTACAGAGTCTGAGCGGGCTGCTGAGCCAGGAGCT 62
QY 50 uLeuProThrAlaLeuLeuValThrArgAlaAlaAlaProGlyArgGlnGln 70
DQ 63 GTTACCAGGAGCTCGCTTGTCTAGTGACTACACGCCATGCCCTACAGGAGGCTGCGAGG 122
QY 70 YArgLeuCysSerProGlnCysAlaGlu-ValArgGlyPheSerAspIysAspIysL 90
DQ 123 CAGATTGTGCTCGCCAGCTGCGCAGACATACCGCGCTTCTCAGACAAAGACAAAAA 182
QY 90 YSTYrPheTYrCysPhePheArgAspGluArgAlaGluArgAlaTYrArgPheVal 110
DQ 193 AGATTCTTCAAGTCTTCTCGGACGAGAGGAGGAGGAGCGCTTACCGCTTCTGTA 242
QY 110 YSGLuAsnGluThrLeuPheAlaLeuCysPheValProPheValCysTrpIleValCys 130
DQ 243 AAGAGACGAGAGCTGTTTGCCTGTGCTTGTGCGCTTGTGCTGCTGCTGCTGCTGCTG 302
QY 130 hrValLeuArgGlnGlnLeuGlyArgAspLeuSerArgThrSerIysThrThr 150
DQ 303 CAGTGTGCGACGACGCTGAGAGCTGCGCGCGGATCTCTCTGATCTCTCAAGACCACTA 362
QY 150 hrSerValTYrLeuPhePheLeuThrSerValLeuSerSerAlaProValAlaAspGlyP 170
DQ 363 CTCTCTGTGTACCTGCTCTTATACACGAGTGTGTAAGTCTGAGGCTGAGGCTGAGG 419
QY 170 roArgLeuGlnGlyAspLeuArgAsnLeuCysArgLeuAlaArgGluGlyValLeuGly 199
DQ 420 CCGGCTTACGAGGAGCTGAGAACGCTGTGTGCTGCTGCGCTGCGCAGGAGGCTCTGATC 479
QY 190 --ArgArgAlaGlnPheAlaGlnGlyLeuGlnLeuGlnLeuGlnLeuGlnLeuGln 209
DQ 490 ATCATAAGACACATCTCAGAGAGGACCTGGAGAAATTCGAAGCTTCCGGGTTCGCAAG 539
QY 209 alGlnThrLeuPheLeuSerCysGlnLeuProGlyValLeuGlnThrGluValThr 229
DQ 540 TTCAACAAATATTTCTCAACAAAGAGATACACGAGGTGCTTAAAAACTGAGGTCACCT 599
QY 229 Yr 229
DQ 600 AC 601

```

RESULT 7

BG965959
 LOCUS 832 bp mRNA linear EST 12-JUN-2001
 DEFINITION 602829753F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4984429 5', mRNA sequence.

ACCESSION BG965959

VERSION BG965959.1 GI:14353595

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 832)
 AUTHORS Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgaps@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10990 row: j column: 14
 High quality sequence stop: 676.

FEATURES

source
 1..832
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4984429"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Co24"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 1.6 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 225 a 211 c 243 g 153 t
 ORIGIN

Alignment Scores:

Pred. No.: 7,42e-60 Length: 832
 Score: 731.50 Matches: 174
 Percent Similarity: 74.30% Conservativity: 37
 Best Local Similarity: 61.21% Mismatches: 59
 Query Match: 22.73% Indels: 15
 DB: 12 Gaps: 5

US-10-029-347-2 (1-625) x BG965959 (1-832)

```

QY 146 SerIysThrThrThrSerValTYrLeuLeuPheLeuThrSerValLeuSerSerAlaPro 165
DQ 1 TCCAAGACCACTACTTCTGTGTACCTGCTCTTCTATCAACAGCATCTCAAGTCTCGAGC 60
QY 166 ValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCysArgLeuAlaArgGlu 185
DQ 61 ---ACCAATGACCCCGGTTCCAGAGAGCTGAGAACGCTGTGTGCTGCGCCAGGAG 117
QY 186 GlyValLeuGly---ArgArgAlaGlnPheAlaGlnGlyGluLeuGlnLeuGluLeu 204
DQ 118 GGCATCTCGATCATCATCAAAAGACAAATTCAGAAAGAGGACCTCGAGAAATTTGAAGCTT 177
QY 205 ArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuProGlyValLeuGlu 224
DQ 178 CCGGTTTCCCAAGTTTCAGACAATATTTCTCAACAAAGAGATACCAAGGAGTGTCTAAA 237
QY 225 ThrGluValThrTYrTYrPheIleAspGlnSerPheGlnGluPheLeuAlaAlaLeuSer 244
DQ 239 ACTGAGGTCACTTACAGTTCATTGACAGAGCTTCCAGGAGTCTTGGCTGCATTGTCA 297
QY 245 TYrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyValGlyThrLeuLeu 264
DQ 298 TACCTACTAGAGCTGAGCGAACTCCGGGAGCGCCCGCAGCGGCTGTCAGAGGCTCTCTG 357
QY 265 ArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeu 284
DQ 358 AACTCTGACGGGAGCTACGTGTGTCTTGCACCTGACCAACCCGATTCCTCTTTGGCTGT 417
QY 285 LeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMetValSerGluArg 304
DQ 418 CTAAATACAGAGGGCTTCGTGACATTCGAAACCAATTTTGGCTGTGTGTGCCAGATCAT 477
QY 305 ValLysGln-GluAlaLeuArgTrpValGlnGlyGlnGly-----GlnGlyCysProGln 322
DQ 478 GTGAACAAGACACCCCTCGGTGGTTCAGAGGACAGACCCACCCAGGGGCCACCACT 537
QY 322 yValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluProGluGlu 342
DQ 538 AGGGCAAAAAAAGACTGTCTGAG-----CTGGAGGACATCGAGGACGCGAGGGA 585
QY 342 uGluGluGluGluGluGluPro---AsnTYrProLeuGluLeuTYrCysLeuTYrGlu 361

```

D6	586	GGAGGAGGAGGAGAGGAGGACCTCAACTCGGACTCGAGCTGCTGTACCGCTGTATGA	645
QY	361	uThrGlnGluASPAlaPheValArgGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuG1	381
DB	646	GACACAGAGGAGGACGG-ATCAGTCGACAGGGCTCTCAGCAGCGCTTCAGAGATAGTACTGGA	754
QY	381	nArgVal-ArgPheCysArgMetAsp-ValAlaValLeuSerTyrCysValArgCysCys	400
DB	705	GCGAGTCCAGGTTGAGCGCGATGGAACCTAGAACGCTCTGAATTACGCATGCCAGCATGC	764
QY	401	ProAla-GlyGlnAla-LeuArgLeuLeuLeuSerCysArgLeuValAlaAlaGlnGluLysL	420
DB	765	CCAGATGGCGGAGGCTCGACCACTGCTGACCTCGGCGAGATCGCGCAAGAGAAAGAA	824
QY	420	YS 420	
DB	825	AA 826	
RESULT 8			
LOCUS	AF054176	2218 bp mRNA linear HTC 22-MAY-2001	
DEFINITION	Homo sapiens angiotensin/vasopressin receptor AII/AVP mRNA,		
ACCESSION	AF054176	complete cds.	
VERSION	AF054176.1	GI:3341995	
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2218) Mao, M., Fu, G., Wu, J.-S., Zhang, Q.-H., Zhou, J., Kan, L.-X., Huang, Q.-H., He, K.-L., Gu, B.-W., Han, Z.-G., Shen, Y., Gu, J., Yu, Y.-P., Xu, S.-H., Wang, Y., Chen, S.-J. and Chen, Z., Identification of genes expressed in human CD34(+) hematopoietic stem/progenitor cells by expressed sequence tags and efficient full-length cDNA cloning Proc. Natl. Acad. Sci. U.S.A. 95 (14), 8175-8180 (1998) 98318631 9853160		
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			

[illegible]

```

375 PheProGluLeuAlaLeuGlnArgValArgPheCysArgMetAspValAlaValLeuSer 394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
773 TTCCCAAGATTGAGATCAAT-----CTCTCCACCAAGATGACCATGTTCTCTCC 826

395 TyrCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuSerCysArgLeuVal 414
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
927 TTTTGCAATGAGAAGCTGCTATCGGTGGAGTCACTGCTCCCTGGGTTTCTCCATAACATG 886

415 AlaAlaGlnGluLysLysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGly 434
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
887 CCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 931

435 GlySerSerGlnGlyThrThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGln 454
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
932 -----ATGGTGCAG 940

455 AlaMetThrAspProLeuCysHisLeuSer---SerLeuThrLeuSerHisCysLysLeu 473
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
941 TGTGTCTCTCCCAAGCTCTCTCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000

474 ProAspAlaValCysArgAspLeuSerGluAlaLeuArgAlaProAlaLeuThrGlu 493
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1001 TCCCATGAGTCTGCTCGACATCTCTTGGTCTGACGACCAACAGAGCTGGTGGAG 1060

494 LeuGlyLeuLeuHisAspArgLeuSerGluAlaGlyLeuArgMetLeuSerGlyGlyLeu 513
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1061 CTGGACCTGAGTCAACAGCCCTCGGTGACTTCCGATCAGACTCTGTGTGGGACTG 1120

514 AlaTyrProGlnCysArgVal----- 520

1121 AAGCACCTGTGTGCAATCTGAAGAAGCTCGTGTGGTGAATCTTGCTTACGTCATCT 1180

521 -----GlnThrValArgVal 525
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

1181 GTTGTTCAGCTTCTGCTCGTACTCAGCACTAATCAGATCTCAGCACCTTACTCGG 1240

526 GlnLeuProAspProGlnArgGlyLeuGlnTyrLeuVal---GlyMetLeuArgGlnSer 544
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1241 AGGCACACTCTCGAGACAGGAGATCAAACTACTCTGTGAGGAGCTCTTGACCCCTGAC 1300

545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMet----- 561
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1301 TGCAAGCTTCAGGTGTGGGTTAGAACACTGCAACCTCAGTCACACTCTCTCTGGGAT 1360

561 ----- 561
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

1361 CTTTCCACACTCTGACCTCAGCCAGAGGCTCGGAAGCTGAGCTGGCCACATGAC 1420

562 -----ValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeu 576
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1421 CTGGCGGACCTGGGGGTGATGATGTTCTGTGAAGTGTGAAACAGCAGAGCTGCTCTG 1480

577 GlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnLeuGlnAla 596
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1481 CAGAACCTGGGGTGTCTGAAATGATTTCAATATGAGACAAAAGTGGCTTAGAACA 1540

597 ValLysArgAlaLysProAspLeuValLeu 606
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1541 CTTCAAGAGAGAAAGCTGACCTGACCGTC 1570

RESULT 9
BY262618
LOCUS
DEFINITION BY262618 RIKEN full-length enriched, visual cortex Mus musculus
cDNA clone K33014H14 5', mRNA sequence.
ACCESSION BY262618
VERSION BY262618.1 GI:26444130
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

REFERENCE

1 (bases 1 to 416)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Boso, H., Kondo, S.,
Niikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Kogami, A., Schombach, C.,
Gotohori, T., Baidarelli, R., Hill, D.P., Buitt, C., Hume, D.A.,
Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K.W., Blake, J.A., Bradt, D., Brusica, V., Chothia, C., Corbani,
L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,
A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
Carvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,
P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,
Ranaachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,
B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C.A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,
R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, C., Aizawa, K.,
Arakawa, I., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-22 Saitama-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-5222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp

URL: http://genome.gsc.riken.go.jp/

Alzawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane

T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,

Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,

Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,

Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct

Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Michela Fagioli and Takao K. Hensch (

Laboratory for Neuronal Circuit Development Brain Science Institute

RIKEN 2-1 Hirotsawa-Wako-shi, Saitama 351-0198 Japan) whose

assistance we gratefully acknowledge. Please visit our web site

(http://genome.gsc.riken.go.jp) for further details.

Location/Qualifiers

1..416

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

FEATURES

SOURCE

Fax: 81-45-503-9216
E-mail: genome-res@genome.gsc.riken.go.jp,
[URL](http://genome.gsc.riken.go.jp/): http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsubashi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kaiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,K., Kawajiri,Y., Shibata,K. and Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details
 e mouse tissues.

	Location/Qualifiers
1..670	/location="Mus musculus"
	/organism="Mus musculus"
	/mol_type="mRNA"
	/db_xref="taxon:10090"
	/clone="A330055K17"
	/sex="male"
	/tissue_type="spinal cord"
	/dev_stage="adult"
	/lab_host="DH10B"
	/clone_lib="RIKEN full-length enriched, adult male spinal cord"
	/note="Site 1: SalI; Site 2: BamHI: cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGACAGAGAGATCCAAAGACTCTTTTTTTTNN 3'] . cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGACAGAGATTCTCGAGTAATAATTAATCCCCCCCCCCC 3'] . cDNA was cleaved with XhoI and BamHI. Vector: a modified plasmid KS(+) after bulk excision from Lambda FLC I."
143 a	189 c 210 g 128 t

ores:

Length:	670
Matches:	121
Conservative:	6
Mismatches:	19
Indels:	1
Gaps:	0

7-2 (1-625) x B8633515 (1-670)

1 MetLeuLaGlnProGln:ArgLeuLeuPheLeuASGLYalaSpGluLeuProAla 20
 ::::
 ATACTGGCGCAACCAATTCGCCTACTGTTCATCCTCGAGCGTGACACGAGCTGCTACT 292

K., K., Itoh
 ., NAs to
 new
 ra, E.,
 Matsuura
 . and
 rmat
 ncome Res.
 , Sugahara
 ion of a
 (2001)
 I., Aizawa
 d
 ed with:
) for
 le spinal
 ' was
 dia
 Riken
 atory in
 in Riken
 | CDNA was
 CDNA was
 erse
 -length by
 alization
 second
 of
 CCCCCCCCCCCC
 r: a
 rom Lambda


```

Qy 21 LeuGlyG.yProGluAlaAlaProCysThrAspProPheGluAlaAlaLaserGlyAlaArg 40
D 293 TTACCATCCTCAGAGGCCACTCCTGCAAGAGCCCTTTGAGGCTACAGTGCCTTGAGA 352
Qy 41 ValLeuGlyG.yLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 60
D 353 GTGCTGAGCGGCTGCTGAGCAGGAGCTGTACCAGGAGCTGCCTTGCTAGTACTACA 412
Qy 61 ArgAlaAlaAlaProGlyArgGluGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
D 413 CGCATCCCGCTACAGGAGCTCGAGGAGAGATGTGCTCGCACAGTGCAGAGAATA 472
Qy 81 ArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLysLys 100
D 473 CGCGCTTCTCAGACAAGACAAAAGATTTCTCAAGTTCCTCCGAGCAGAGAGG 532
Qy 101 ArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
D 533 AAGCAGAGCGCGCTACCGCTTCCGTAAGAAGA-AACGAGAGCGCTGTTGCACTGCTTT 591
Qy 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluGlyArg 140
D 592 GTGCCCTTCGGTCTCGATCGTGTGCACAGTGTGCAGCAGCAGCTAGAGCTGGGCGCG 651
Qy 141 AspLeuSerArgThrSer 146
D 652 GGTCTCTCTCGTACATCC 659

RESULT 1:
LOCUS BC021272 3368 bp mRNA linear HTC 18-JAN-2002
DEFINITION Homo sapiens, c-one IMAGE:4811303, mRNA.
ACCESSION BC021272
VERSION BC021272.1 GI:1820422
KEYWORDS HTC
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg,R.
Direct Submission
Submitted (14-JAN-2002): National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2530,
USA
NTH-YGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcaps-remail.nh.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
Lim,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C.,
McMowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
Tongson,E.E., Touchman,J.W., Tsurgon,C., Vogt,J.L., Walker,M.A.,
Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IMAGE Plate: 39 Row: 9 Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Genomescan gene prediction
This clone has the following problem: frame shifted.

```

```

FEATURES
    source          Location/Qualifiers
    1..3368         /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:4811303"
                  /tissue_type="Placenta, choriocarcinoma"
                  /clone_lib="NTH MGC 21"
                  /lab_host="DH10B-R"
                  /note="vector: pOTB7"
BASE COUNT      900 a 803 c 828 g 837 t
ORIGIN
Alignment Scores:
Pred. No.:      2..89e-46      Length:      3368
Score:          598.50        Matches:    193
Percent Similarity: 44.56%    Conservat: 114
Best Local Similarity: 28.01%  Mismatches: 251
Query Match:    18.60%       Indels:     131
DB:             11           Gaps:       19
US-10-029-347-2 (1-625) x BC021272 (1-3368)
Qy 1 MetLeuAlaGlnProGlnArgLeuPheIleLeuAspGlyAlaAspGluLeuPro--- 19
D 725 ATCGTGTCTCAACCGGAGAGACTCTTGTTCGTCATCGACAGCTTCGAAGAGCTCGAGGC 784
Qy 20 AlaLeuGlyG.yProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAla 39
D 785 GCCTTGAACGACCCGATTCGATCTGTGTGACTGTATGGAGAAACGCCGCGTGCGAG 844
Qy 40 ArgValLeuGlyG.yLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 59
D 845 GTGCCTCTGAGCAGTTTCTGAGGAAGAAGATGTCTCCGAGGAGCTCTCCTCATCCGC 904
Qy 60 ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu 79
D 905 ATCAACCCGCTGTGCCCGAAGAGCTCCGGATCAGTGACGATCTCAGAAATCTTACCAG 964
Qy 80 ValArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLys 99
D 965 CCCCGGGATTCACAGCAGATGATAGTTAGTTATTTCTGCTGTTCTTCAAGACCCG 1024
Qy 100 ArgArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCys 119
D 1025 AAAAGAGCCATCGAAGCTTCAATCTTGTAAAGAAAGTGAACAGCTGTTTCCATATGC 1084
Qy 120 PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGly 139
D 1085 CAAATCCCGCTCTCTGCTGATCTGTGTACCATCTTGAAGCAAGAGATCGAGAAAGGA 1144
Qy 140 ArgAspLeuSerArgThrSerLysThrThrSerValThrLeuLeuPheIleThrSer 159
D 1145 AAAGACCTGGCCCTGACCTGCCAGAGCACTACCTCTGTGTACTCTCTTCTGCTTTAAC 1204
Qy 160 ValLeuSerSerAlaProValAlaAspGly-----ProArgLeuGlnGlyAspLeuArg 177
D 1205 ---CTGTTCAACACCTGAGGTCGCCGAGGCCGACTCCGAAACCCAGCAGCTGAAG 1261
Qy 178 AsnLeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLys 197
D 1262 GCCCTGTGCTCTCCCTGGCTGACAGGGGTATGTGGACAGACACATTTGAGTTTGTGAAGAC 1321
Qy 198 GluLeuGluGlnLeuLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLys 217
D 1322 GACCTCCGAGAAATGGGGTTGTGACGCTACATCCCTCGCTG---CTGGGACCAAG 1378
Qy 218 GluLeuProGlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGln 237
D 1379 ATACTTCTGAAGTACGGGAGCGTGAGAGCTCTACGTGTTCTCTCCACGCTGTGTATCCAG 1438
Qy 238 GluPheAlaAlaLeuSerTyrLeuLeuGluAspGly---GlyValProArgThrAla 256
D 2388 GGTCTCTCTCGTACATCC 2599

```

```

Db 1439 GAGTCTGTGCGCGCTGCTCTATTGCTCAGAGCCACCTTGATCATCTCCACCCAGCT 1498
Qy 257 AlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnPro-----HisSerHisLeu 274
Db 1499 GTGAGATGTGTACAGAAATGCTAGTCCCAATTTTGAAAGAGCAAGAGAGACATGG 1558
Qy 275 ValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluAsgMetArgAspLeuGlu 294
Db 1559 ATTTTGGGTGGTGTCTTCTACTGGCTTTTAAATAAAGAGAACAAAGAAAACTGGAT 1618
Qy 295 ArgHisPheGlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTrpValGln 314
Db 1619 GCGTTTTTGGCTTCCAACTGTCCCAAGAGATAAAGCAG----- 1657
Qy 315 GlyGlnGlyGlnGlyCysProGlyValAlaProGluValThrGluGlyAlaLysGlyLeu 334
Db 1658 -----CAAATTCACCACTGGCTGAAGAGCTTA 1684
Qy 335 GluAspThrGluProGluGluGluGluGlyGluGluProAsnTyProLeuGlu 354
Db 1685 GGGGAGCGTGGCAATCCTCAGGACAGAGTGGATTCC-----TTGCGG 1726
Qy 355 LeuLeuTyCysLeuTyGluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArg 374
Db 1727 AATTATTACTGTCTTTGAAATGCAGGATCCTGCTTTGCGAAGCAGCAGT----- 1780
Qy 375 PheProGluLeuAlaLeuGlnArgValArgPhe-----CysArgMetAspValAla 391
Db 1781 -----ACCTCTCTCAGAAAGCTAACTTTTCATATTATGACCAAGCTGGACTTGTG 1831
Qy 392 ValLeuSerTyCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuIleSerCys 411
Db 1832 GTTCTGCTACTGCTTAAATACTGCTCCAGCTTGAGGAAACTC---TGTCTTCCGTT 1898
Qy 412 ArgLeuValAlaAlaGlnGlyLysLysLysLysSerLeuGlyLys----- 426
Db 1889 CAAATATGCTTTAAGAAAGAGGATGAACACAGCTCTACGTCGGATTACAGCCTCATCTGT 1948
Qy 426 ----- 426
Db 1949 TGGCATACATCTGCTCTGTGTCACCAACCGGCGACCTCAGAGAGCTCCAGTGCAG 2008
Qy 426 ----- 426
Db 2009 GACAGCACCTCAGCGAGTCGACCTTTGTGACCTGGTGTAAACCACTGAGGCATCCACG 2069
Qy 427 ---ArgLeuGln-----AlaSerLeuGlyGlyLys----- 436
Db 2069 TGTGCGCTTCAGAAGCTTGAATAAATAACGTTTCTTTTCGGCCAGAGTGTCTGCTC 2126
Qy 437 -----SerGlnGlyThrLysGlnLeuProAlaSerLeu----- 448
Db 2129 TTTGAGGTGCTTTTATCAGCCAGACTTGAATACTCCTGAGCTTCACCTCAGCAAACTC 2189
Qy 445 -----LeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSer 464
Db 2189 TCTCGTGATGACATCAGTTCCTCTGTGTGCTGTGAACCTCCAGCAGGCAACGTCAA 2248
Qy 465 SerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerGluAla 484
Db 2249 GAGCTAGCGCTGGTAAATGTGCACCTCTCACCCATTTGATTGTGAAGTCTCTGCTGGCCTT 2308
Qy 485 LeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAla 504
Db 2309 CTAACCAACAAACAGAGCTGACGTATCTGAATGTATCTCTGCAACCAAGTTA---GACACA 2365
Qy 505 GlyLeuArgMetLeuSerGluGlyLeuAlaTpproGlnCysArgValGlnThrValArg 524
Db 2365 GGGGTGCCCCCTTTGTGTGAAGCCCTG-----TGCAGCCCAACAGACGCTGCTG 2413
Qy 525 ValGlnLeu-----ProAspProGlnArgGlyLeuGlnTyLeuValGly 539
Db 2414 GTATACCTGTATGTGGCTTTTGTGCCACCTCAGCGAGCAGTGTCTGGCAATACATCTCTGAA 2473

```

```

Qy 540 MetLeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAla 559
Db 2474 ATGCTTCTCGTAAACAGAGCGTGCCTATCTAGACCTCAGTGCCCAATGTCTCTGAAGGAC 2533
Qy 560 ProMetValThrTyLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeu 579
Db 2534 GAAGGACTGAAACTCTCTGCGAGGCTTGAACAATCCGAGACTGCTGCGCTGGATTCACTG 2593
Qy 580 -----SerLeuAlaSerValGluLeu 586
Db 2594 TGTGTTGTAATAATGTTTATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2653
Qy 587 SerGluGlnSerLeuGlnGluLeuGln 595
Db 2654 AGCAATCAAAACCTGAAGATTCTGCAA 2680

RESULT 12
BC012789 3369 bp mRNA Linear HTC 04-MAR-2003
LOCUS
DEFINITION Homo sapiens, clone IMAGE:3957172, mRNA.
ACCESSION BC012789
VERSION BC012789.1 GI:15215377
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3369)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (15-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs@email.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LJNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettenan, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LJNL at: http://image.llnl.gov
Series: IRAL Plate: 25 Row: k Column: 23
This clone has the following problem: retained intron.

FEATURES
source
1..3369
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3957172"
/tissue_type="Placenta, choriocarcinoma"
/clone_id="NIH_MGC_21"
/lab_host="DH10B-R"
/notes="vector: pOT8"

BASE COUNT 901 a 803 c 828 g 837 t

ORIGIN
Alignment Scores: 2.88e-46 Length: 3369
Pred. No.: 598.50 Matches: 193
Score: 44.56% Conservative: 114
Percent Similarity: 28.01% Mismatches: 251
Best Local Similarity: 18.60% Indels: 13
Query Match: 11
DB:

US-10-029-347-2 (1-625) x BC012789 (1-3369)

```

1	MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeuPro---	13
725	ATCGTGTCTCAACCGGAGAGACTCTTGTTCGTATGCACAGCTTCGAAGAGCTGCAGGCG	784
20	AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAla	39
785	GGCTTGAACCAACCCGATTCGGATCTGTGTGGTGACTTGTATGTAGAGAAACCGCGGTGCAG	844
40	ArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr	59
845	GTGCTTCTGACGAGTTTGTCTGAGAAAGAGATGCTCCGGAGGCTCCCTGTCTCATCGCC	904
60	ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu	79
905	ATCAAAACCCGTGTGCCGGAAGAGCTCCGGGATCAGGTGACGATCTCAGAAATCTACCAG	964
80	ValAtsGlyPheSerAspLysLysLysTyrrPheTyrrLysPheArgAspGlu	99
965	CCCCGGGATTCAACGAGACGATAGTTAGTGTATTTCTGCTGTTTCTTCAAGACCCG	1024
100	ArgArgAlaGluArgAlaTyrrArgPheValLysGluAsnGluThrLeuPheAlaLeuCys	119
1025	AAAAGAGCCATGCAAGCCTTCAATCTTGAAGAGAAAGTCAACAGCTGTTTCCATATGC	1084
120	PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGly	139
1085	CAAAATCCCCTCTCTGCTGTGATCTCTGTGTACCAGTCTGAAGCAAGAGATGCAAGAAGGA	1144
140	ArgAspLeuSerArgThrSerLysThrThrSerValTyrrLeuLeuPheIleThrSer	159
1145	AAAAGACTCGCCCTGACCTCCAGAGACACTACCTCTGTACTCTCTTTCTGCTTTAC	1204
160	ValLeuSerSerAlaProValAlaAspGly-----ProArgLeuGlnGlyAspLeuArg	177
1205	---CTGTTTCAACCTCAGGGTGCAGGGGCCGACCTCCGAAACCCAGCACCACTGCTGAG	1261
178	AsnLeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLys	197
1262	GCCTGTGCTCCCTGCTGCAGAGGGTATGTGGACAGACACATTGTAGTTTGTGAGAC	1321
198	GluLeuGluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLys	217
1322	GACCTCCGGAGAAATGGGTTGTGTGACGTGCATCCCTCGCTG---CTGGGCACCAAG	1378
218	GluLeuProGlyValLeuGluThrGluValThrTyrrGlnPheIleAspGlnSerPheGln	237
1379	ATACCTTCTGAAGTACGGGAGCCGTGAGAGCTCCTAGTGTTCCTCCACGTGTGTATCCAG	1438
238	GluPheLeuAlaAlaLeuSerTyrrLeuLeuGluAspGly---GlyValProArgThrAla	256
1439	GAGTTCTGTGCGCCCTTGTTCTATTTTGCTCAAGAGCCACCTTCATATCTCTACCAGCT	1498
257	AlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnPro-----HisSerHisLeu	274
1499	GTGAGATGTGTACAGAAATTCCTAGTTGCCAATTTTGAATAAAGCAAGGACAGCATTGG	1558
275	ValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGlu	294
1559	ATTTTTTGGGGTGTTTTCTTAACCTGGCCTTTTAAATAAAGAACCAAGAAAAAATCGAT	1618
295	ArgHisPheGlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTrpValGln	314
1619	CGCTTTTGGCTTCCAACTGTCTCCCAAGACATAAGCAG-----	1657
315	GlyGlnGlyGlnGlyCysProGlyValAlaProGluValThrGluGlyAlaLysGlyLeu	334
1658	-----CAAATTCACCAGTCCCTGAAGAGCTTA	1684
335	GluAspThrGluGluProGluGluGluGluGlyGluGluProAsnTyrrProLeuGlu	354
1685	GGGAGCGTGGCAATCTCTCAGGACAGGTGGATTC-----TTGGCG	1726

RESULT : 3

AK054264

Accession	LOCUS	AKC54264	3400 bp	mRNA	linear	HTC 05-DEC-2002
AK054204						

DEFINITION	Mus musculus 2 days pregnant	adult female ovary	CDNA	RIKEN
LOCUS	AK054264	3400 bp	1118nt	H19 03-DE

DEFINITION Mus musculus 2 days pregnant adult female ovary cDNA, K12EN f111-length enriched library clone:E330007A02 product:weakly

FAST-RENGEN ENRICHED LIBRARY, CLONE:E33000/A02 product:WEAKLY

similar to PAN1, full insert sequence.

AK054264
AK054264.1 GI:26344106

Hit; CAP trapper.

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCES

Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
0349636

REFERENCE

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

REFERENCE

Shibata, K., Itoh, M., Aizawa, K., Nagacka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishibe, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed-sequence
Genome Res. 10 (11), 1757-1771 (2000)
2350913
11076861

REFERENCE

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kado, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, A., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsi, G., Blake, J., Boffelli, D., Sojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schenbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kottuk, S.,
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

REFERENCE

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3400)

REFERENCE

Adachi, J., Aizawa, K., Akimura, T., Hara, A., Hashizume, W.,
Fukuda, S., Furuno, M., Haragaki, T., Hiraoka, K., Hiraoka, T., Hirozane, T.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Kikuchi, T., Kikuchi, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Chisato, N.,

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tanaka, M., Tagawa, A., Takahashi, F., Takaku-Akai, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
Direct Submission

TITLE

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES

source

1..3400
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:E330007A02"
/db_xref="taxon:10090"
/clone="E330007A02"
/sex="female"
/tissue type="ovary"
/clone lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days pregnant adult"
76..2553
/note="unnamed protein product; Weakly similar to PAN1
(SCOP152047, evidence: SCOP)
putative"
/codon_start=1
/protein_id="BAC35710.1"
/db_xref="GI:26344107"
/translation="MEHFDPLGFLNKLVDLTKTELNNFKRLSCSLPTVYKQINK
LTMLANGAQLAEILTDHCPASWIKRVTVQILEENRVDLAEVLVQIEAEVLVPEE
KYSSPREPSGTLTPFNNFVQAGKPEDEKEMKRYTAKMKNFPCNKKEIYVVT
ESYKTLALCNPKITPFAHAIIVLHGPQSGKTTWAKOLMEWSESKQAQIFSCAFYI
SCRYNNPKCTFAHLSDMNPSCDVCVIRDLIGKEFLFVVDGDFELTFPAGALRO
LCGDNWTKPVEVLGSLKKRMMAHAILVTTRTQSLHQLFVMDQPLVLTGLGLE
QSKQYFQYFDEEEDEEKGKALRKALKEVRNADLYOMASLPTACIGFCLCLE
RMKGEDLSLTQYTSNPLNFCVFSSETCHLNEEFOLFKKICILAAISLLEQ
VTILCEEDFLTLKLNMLNLPVCRHILFKDSSSTHCLSFICLGIQCLAAIFVQEL
GOESKGVSKYSIONLSREARLNKLPDGLLPFVFLNNEIRIOLKTLTFGQISTEV
KRFLECSGKGNKPLLLMNMORILSCLYBSDEEGFVKEMVLFEDISLHKLSTDL-
HASFCLKNSQNLQWLSLKVKAAPPENVAALSTAKHORSPOEQLMFTWTFDCTFN
SNKLVLFD:HESFLNSALEILCEKLPASCCLOKVLKINISPPDAYEKCLLFNGY
KTSIHILQGGNLSMHSLCEVLPKACNKLFLSGSCSTAAQKWDFFFPVLKYNQS
LIFLDLTONSLDKSAKLLCNWKEPK"

CDS

3383..3388
/note="putative"
3400
/note="putative"

polyA_signal

polyA_site

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-029-347-2 (1-625) x AK054264 (1-3400)

QY

6 GlnArgLeuLeuPheLeuLeuAspGlyAlaAspGluLeu-----ProAlaLeuGlyGly 23

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-029-347-2 (1-625) x AK054264 (1-3400)

QY

6 GlnArgLeuLeuPheLeuLeuAspGlyAlaAspGluLeu-----ProAlaLeuGlyGly 23

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-029-347-2 (1-625) x AK054264 (1-3400)

QY

6 GlnArgLeuLeuPheLeuLeuAspGlyAlaAspGluLeu-----ProAlaLeuGlyGly 23

```
Db      835  AAGGAATTCCTATTGTGGTAGATGGTTTGTGATGAGCTGACATTCACGCA---GGAGCT 891
Qy      24   ProGluAlaProCysThrAsp-----ProPheGluAlaIAspGly 38
Db      892  CTGATCCGTGACCTTTGTGGTGAATCGGAACACAGTCGAAGCCTGTAGAGGTC----- 942
Qy      39   AlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuVal 58
Db      943  -----TACTGGCAGATTTCCTGAAGAGGAGATGGCACTATGCCACCCTGCTGGTA 996
Qy      59   ThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla 78
Db      997  ACCACAGGCACACAGTCTTTSCACCAATCTTTGTCATGATGGATCAGCCACTACTAGTA 1056
Qy      79   GluValArgGlyPheSerAspLysAspLysLysLysLysPheTyLysPhePheArgAsp 98
Db      1057  GAAACTCTGGGCTTCTTGGAGACAAAGAGCAGGATCTTTCAGAAATATCTTGAAGAT 1116
Qy      99   GluArg-----ArgAlaGluArgAlaTyArgPheVal 109
Db      1117  GAGAGGGTGAGGAGGAGGAATAAGGTGAGGGAAGGCATCTAGAGCTCTCAAGGAAGTG 1176
Qy      110   LysGluAsnGluThrLeuPheAlaLeuCysPheValProPheValCysTrpIleValCys 129
Db      1177  AGATGCAATGCTGATCTATACAAAATGCCCTCACTTCTCTACTGCATGTGGGATTTCTGC 1236
Qy      130   ThrValLeuArgGlnGlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThr 149
Db      1237  CTCTGTCTGGAGCTTAAGATTAAGAAGGGGAAGACCTGTCTCTGACATGTGACATAC 1296
Qy      150   ThrSerValTyLeuLeuPheIleThrSerValLeuSerSerAlaProValAlaAsp--- 168
Db      1297  ACCTCTATGTTCTTGAACTTCTGTGGAGAGTCTTCTCATCAGAAACCTGTGAGGACCAT 1356
Qy      169   --GlyProArgLeuGlnGlyAspLeuArgAsnLeuCysArgLeuAlaArgGluGlyVal 187
Db      1357  CTCAATGAGGAATTCCAAATCTTATTCAAGAAATATGTATCTCTGGCTGCGCAATAGTCTC 1416
Qy      166   LeuGlyArgArgAlaGlnPheAlaGlyLysLysLysLysLysLysLysLysLysLys 237
Db      1417  TTGGAACAGGTGCCCACTACTGTGTGAAGAGACTCTTTGACATTAATAATTAACCTAAAC 1476
Qy      208   LysValGlnThrLeu-----PheLeuSerLysLysGluLeuProGlyValLeu 223
Db      1477  AATTGGATCCCATGGTGTGAGGCACATCCTCTTTAAG----- 1515
Qy      224   GluThrGluValThrTy-----GlnPheIleAspGlnSerPheGlnGluPheLeuAla 241
Db      1516  GACAGTTCACAGTACACACTGCTCTCTTTCATCTGCTCGGTATCCAGCAACTCCTGGCT 1575
Qy      242   AlaLeuSerTyLeuLeuLeuGluAspGly-----GlyValProArgThrAlaAla 257
Db      1576  GCAATAATATTGTTCAGAGCTTGGACAGAAAGCAAGGGGTGTCTCTAAATATAGT--- 1632
Qy      258   GlyGlyValGlyThrLeuLeuArgGlyAspAlaGln---ProHisSerHisLeuValLeu 276
Db      1633  -----ATACAGAACATGCTTTCAGGGAAGCAGAGATTAAAAACCCCTGACCTGTCTGGG 1686
Qy      277   ThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHis 296
Db      1687  CTGCTGGCAATTGTATTGGCTCTCTGAATGAGACTCGTATTCAAGAGATTGAAGACTACT 1746
Qy      297   PheGlyCysMetValSerGluArgValLysGlnAlaLeuArgTrpValGlnGlyGln 316
Db      1747  TTGGCTGCCAAATATCAACTGAGGTCAAGAGAAAAATTCCTG----- 1788
Qy      317   GlyGlnGlyCysProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAsp 336
Db      1788  ----- 1788
Qy      337   ThrGluGluProGluGluGluGluGluGluGluGluGluGluProAsnTyProLeu----- 353
```

```
Db      1789  -----GAGTGTGAGTCAGGGGA-----AATAAACCCCTTACTATTACTG 1827
Qy      354  -----GluLeuLeuTyLysLeuTyLysLeuThrGlnGluAspAlaPheValArg 369
Db      1828  ATGAATATGCAAGAAATCTTGTCTGTCTCTATGAATCTCAGGAGGAGGGTTTGTGAAG 1887
Qy      370   GlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPheCysArgMetAsp 389
Db      1888  GAAGCATATGGTCTCTTTGAGGACATCTCTCTTGCAATTTGAAA-----ACCAGCACTGAC 1941
Qy      390   ValAlaValLeuSerTyCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuIle 409
Db      1942  TCATTTATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1992
Qy      410   SerCysArgLeu-----ValAlaAlaGlnGluLysLysLys 421
Db      1993  TCCTCTGAAAGTAGAAAGGCGTCTTTCCAGAGAATGTGTGTCACACTAGTCCACAGCT 2052
Qy      422   Lys----- 422
Db      2053  AAGCATCAAGGTCCCGATGAACAACGCATGTTAACTTTCTGGACAGACTTTTGTGTAC 2112
Qy      423  -----SerLeuGlyLysArgLeu-----GlnAlaSerLeuGlyGly 434
Db      2113  ACATTTAATTCAAATAAGAACTGTGTATTTCTAGACATCCATGAAAGCTTCTCAACAGC 2172
Qy      435   GlySerSerGlnGlyThrThrLysGlnLeuProAla----- 446
Db      2173  TCTGCACTGGAGATTCTTTGTCAAAAGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2232
Qy      446  ----- 446
Db      2233  GTCTCTCAAAATATTTCTCCAGATGATGCTACGACGAAATTTGTGCTCATTTTAAATGGT 2292
Qy      447  -----SerLeuLeuHisPro 451
Db      2293  TATAAAACCATCTCACATCTGATCTGCAAGGTGGCAACCTGGACAGCATGCACTCA 2352
Qy      452   LeuPheGlnAlaMetThrAspProLeuCysHisLeuSerSerLeu----- 466
Db      2353  CTGTGTGAGGTCTTCAAAATCTCAGCTGCAACCTGAAAGTTTCTCAGCCTGGATCTTGT 2412
Qy      466  ----- 466
Db      2413  TCTACTGTGCTCAGAAATGGATGACTCTTCTCCAGTCTTGAAAGTCAAACATCCCTG 2472
Qy      466  ----- 466
Db      2473  ATATTCTGGATCTCACAGACAAACAGTCTCTTGGACAAAAGTGTAAAGTCTCTATGTAA 2532
Qy      467  -----ThrLeuSerHisCysLysLeuPro 474
Db      2533  ATTTGGAAGGAACCAAGTGAATACCTGACAGGGGTATCGTGGAGAACTGTGAGCTCACT 2592
Qy      475   AspAlaValCysArgAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeu 494
Db      2593  GAAGCTCTGTCGAAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2652
Qy      495   GlyLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAla 514
Db      2653  AGCCTGGCCAAACAATAAATCTGGGGATAACGAGGTGAAAAACCTGTGTGAAAGCAATAGC 2712
Qy      515   TrpProGlnCysArgValGlnThr---ValArgValGlnLeuProAspProGlnArgGly 533
Db      2713  TGCACGGAGTGAATCTATACAGACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2772
Qy      534   LeuGlnTyLeuValGlyMetArgGlnSerProAlaLeuThrThrLeuAspLeuSer 553
Db      2773  TGCCACTATCTCTCAAGATGCTCAACAAACACTGACGCTGAAACACTTGGATCTGGGG 2832
Qy      554   GlyCysGlnLeuProAlaProMetValThrTyLeuCysAlaValLeuGlnHisGlnGly 573
Db      2833  CTAAATCGATCGGAACCAAGGAGCAAAAGTTTCTGTGTGAGGCTTTGAAGAACCCCAAG 2892
```

Qy 574 CysGlyLeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnGlu 593
 Db 2893 ACTAAATGAAAAGCTTGGCTGTGTGGATGCTCCATCATCTCCACTGCAAGTGCACAGAC 2952
 Qy 594 LeuGlnAlaVallyArgAlaLysProAspLeu 604
 Db 2953 TTCTCTGAGACTTTAGACGACACAGAGCCGTG 2985

RESULT : 4
 BY281505 431 bp mRNA linear EST 11-DEC-2002
 LOCUS cDNA clone K430354M18 5', mRNA sequence.
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Mus musculus (house mouse)
 Mus musculus

REFERENCE
 AUTHORS

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
 Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 (bases 1 to 431)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaide, T., Ootani, N., Saito, R., Suzuki, H., Yamashita, I., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gajobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
 Guackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisei, K. W., Blake, J. A., Bradt, D., Brusio, V., Chochia, C., Corbani,
 L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
 A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
 Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. W.,
 King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lechard, B., Lyons,
 P. A., Maglott, D. R., Nakats, L., Marchionni, L., McKenzie, L., Niki,
 R., Nagasaita, T., Numata, K., Okido, T., Pavan, W. C., Perce, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, C. C., Reed, D. J., Reid, J., Ring,
 B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou,
 R. D., Tomita, M., Verardo, R., Wagner, L., Wahlested, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,
 M., Yang, Y., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, K., Imotari, K., Ishii,
 Y., Itoh, M., Kagawa, S., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
 K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Jander,
 E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12468851

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0645, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, K., Carninci, P., Fukuda, S., Hirozane,

T., Imotari, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,

Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, N., Sakai, K.,

Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,

M., Waki, K., Wataniki, A., Muramatsu, M. and Hayashizaki, Y. Direct

Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Michela Fagioli and Takao K. Hensch (

Laboratory for Neuronal Circuit Development Brain Science Institute

RIKEN 2-1 Hirasawa, Wako-shi, Saitama 351-0198 Japan) whose

assistance we gratefully acknowledge. Please visit our web site

(http://genome.gsc.riken.go.jp) for further details.

FEATURES

Location/Qualifiers

1..431

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="K430354M18"

/tissue type="visual cortex"

/clone_lib="RIKEN full-length enriched, visual cortex"

BASE COUNT 109 a 113 c 114 g 95 t

ORIGIN

Alignment Scores:

Pred. No.: 1,07e-36 Length: 431

Score: 487.00 Matches: 101

Percent Similarity: 83.33% Conservative: 19

Best Local Similarity: 70.14% Mismatches: 22

Query Match: 15.13% Indels: 2

DB: 13 Gaps: 2

US-10-029-347-2 (1-625) x BY281505 (1-431)

Cy 150 ThrSerValTyrLeuLeuPheThrSerValLeuSerSerAlaProValAlaAspGly 169

Db 2 ACTTGTGTACCTGCTCTTCATCACCAGCATGCTCGAGTCTGCAGGC---ACCAATGGA 58

Cy 170 ProArgLeuGlnGlyAspLeuArgAsnLeuCysArgLeuAlaArgGluGlyValLeuGly 189

Db 59 CCCGGGTTCCAGGAGAGCTGAGACGCTGTGTCTGCTGGCAGGAGGCGATCCTGGAT 118

Cy 190 ---ArgArgAlaGlnPheAlaGluLysGluLeuGluGlnLeuGluLeuArgGlySerLys 208

Db 119 CATCATAAAGACAAATCTTCAGAAAGAGGACCTGGAGAAATTGAAGCTTCGGGTTCCCAA 178

Cy 209 ValGlnThrLeuPheLeuSerLysGluLeuProGlyValLeuGluThrGluValThr 228

Db 179 GTTCAGACATATTTCTCAACAAGAAGAGATACAGAGAGTGTCTGAACCTGAGGTCAAC 238

Cy 229 TyrGlnPheLeuAspGlnSerPheGlnGluPheLeuAlaLeuSerTyrLeuLeuGlu 248

Db 239 TACCAGTTCATTGACCAGAGCTTCCAGGAGTCTTGGCTGCATTGTACATACCTACTAGAA 298

Cy 249 AspGlyGlyValProArgThrAlaAlaGlyValGlyThrLeuLeuArgGlyAspAla 268

Db 299 GCTGAGCAACTCCGGGAGCGCCCGGAGCGGTGTGCAAGCTCTGAACTCTGACCGG 358

Cy 269 GlnProHisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGlu 288

Db 359 GAGCTAGCTGGTCACTCTGCATCTGACCAACCCGATTCCTCTTTGGCTGTCTAAATACAGAG 418

Cy 289 ArgMetArgAsp 292

Db 419 GGGCTTCGTGAC 430

RESULT : 5

AK087843

LOCUS

3577 bp mRNA linear HTC 05-DEC-2002

DEFINITION Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330028A19 product:Similar to pAN2 protein, full insert sequence.

ACCESSION AK087843.1 GI:26352787

VERSION AK087843.1

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20493974

PUBMED 1042159

REFERENCE 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kenno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Chara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20539523

PUBMED 11076861

REFERENCE 4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,C., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamahara,I., Saito,T., Okazaki,Y., Gotohori,T., Bono,H., Kasukawa,T., Saito,R., Kadoya,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Feischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Ma-suo,Y., Nikaudo,I., Pesole,G., Quackenbush,J., Schiraldi,M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,B., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kariya,M., Lee,N.H., Lyons,P., Vaccaroni,L., Mashima,J., Mazzarelli,J., Kombarits,P., Nordone,P., Ring,B., Rongwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,S., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L., Wysshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohtsuki,S. and Hayashizaki,Y. Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 11217851

PUBMED 11217851

REFERENCE 6 (bases 1 to 3577)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,

Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohmoto,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akashira,S., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ Location/Qualifiers

FEATURES

source 1..3577

organism="Mus musculus"

mol_type="mRNA"

strain="C57BL/6J"

db_xref="FANTOM:DB:E330028A19"

db_xref="taxon:10090"

clone="E330028A19"

sex="female"

tissue_type="ovary"

clone_lib="RIKEN full-length enriched mouse cDNA library"

dev_stage="2 days pregnant adult"

309..3257

/note="unnamed protein product; Similar to PAN2 protein (SPTR|Q99MW0, evidence: FASTY, 78.9%ID, 100%length, match=224)

putative"

codon_start=1

protein_id="BAC40024.1"

db_xref="GI:26352788"

translation="MASPFSDFGLMWLYDELNKKEFVKFELKQEVQLGLQKQVST EYKASRODASLLKLVKVEKPAKMTDFEFQKINRDLIERAKEIDGCPKLYEAMH KTKTHDSRAFTISIQNLEKEDTEDDYDCFENLFQSKGTESQPVVPLSGGAGVGK TMLKRLMIAVIESPVFLHKFSYIFCCREVKQLKTASLAELISREWFGSPAPLEEI LSKPEKLLFIIDSLEGMEDLFKWESELCNCTEKQPVNLLSLRKMPLSESLI SATPESFEKMRNIEYTHVKIKGLKERNIKMSFHLFQDRNRAHEAFSLVRENQLF TVCPVLCMWVATCKEIEKGRDPVSICTTSLYTHIFNLFIQNAHSPSKKSQ DOLOGLCSLAEGWMTDTFVGEALRENGIMDSIDIPILLDGLMINTRESEKSVIFL HPSQEVCAAIPLYLLKSHVDHPSQEVKSEIKLMPAFKAKVQW:FGSFIQGLLHES EOKKLEAFHGHQSOEIKRQYQCLTISGNEEIQEQ:DGWKLFCYCFEMODDTLVE ANNCQEVFAKQSDVIAAHKCHGCTFLKLLKLSFTQNVLSGQASHYMERCLTAW NHIQSVFTISKDIQELRMKDTLSSEAFSVLYNNLKHNYTLNVLAVANNVFCVCKYL FFEIQCINQLNLNLSLTILSHSDVLSCLVLSQAEICNIEELVVAACSISDDCKVFA LRVNKTLLHNLASNTLDKGLASLCKSLCHPCDILHLV:ANCSLNEKQMDVISEV LARNKTLSHLDISSNDLDEGLKVCALTIPLDSVLKSLSLRHCLITISGQDLAEVL RNNQLVLSQVSNKLEDTGKLLCDATKHPNCHLEDIGLEACEUTAGACCEDLASTFT QCKTLWAMNLLKALDYNGLVLEALKQQTCAATYVGLGLQITDFDTFOAFLVABQEK NPCLRLSSL"

3557..3562

/note="putative"

3577

polya_signal

polya_site

BASE COUNT 1060 a 728 c 786 g 1003 t

ORIGIN

Alignment Scores:

Pred. No.: 3,73e-34 Length: 3577

Score: 473.00 Matches: 174

Percent Similarity: 39.08% Conservativeness: 116

Best Local Similarity: 23.45% Mismatches: 254

Query Match:	14.70%	Indels:	196
DB:	11	Gaps:	20
US-1.0-029-347-2 (1-625) x AK087843 (1-3577)			
Qy	1	MetLeuAlaGlnProGlnArgLeuLeuPheLeuLeuAspGlyAlaAspGluLeuProAla	20
Ds	960	ATACTATCCAACTCTGAGAACTCTTATTTATCATCGACGCTTAGAAGGATGAAATGT	1019
Qy	21	---LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAla	39
Ds	1020	GATTTCCTCAAAATGGGAATCAGAGCTGTGTGTAATCTGTACAGAGAGCAGCCAGTGAAT	1079
Qy	40	ArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuLeuValThr	59
Ds	1080	GTACTGCTCAGCAGTTTGCTCAGGAGGAAGATGCTCCCTGGAATCCTCTCTCTCATCTCT	1139
Qy	60	ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu	79
Ds	1140	GCTACCCCGCAGAGAGTTTTCAGAAATATGGAGAACAGGATTGAATACACACATGTGAATAA	1199
Qy	80	ValArgGlyPheSerAspLysLysLysLysLysLysLysLysLysLysLysLysLysLys	99
Ds	1200	ATAAAGGACTTAAGAGAGAAATATTAGATGAGTTTCCACAGATTATTCACAGATAGG	1259
Qy	100	ArgArgAlaGluArgAlaThrArgPheValLysGluAsnGluThrLeuPheAlaLeuCys	119
Ds	1260	AACAGAGCCCATGAAGCCCTCAGTCTGGTGAGAGAAATGAGCAGCTGTTCTCCTGTATGT	1319
Qy	120	PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuLeuGly	139
Ds	1320	CAGGTCCCTCTGCTCTCTCTGATGCTGGTGTGCTCTGCTCTCTCTCTCTCTCTCTCTCT	1379
Qy	140	ArgAspLeuSerArgThrSerLysThrThrThrSerValThrLeuLeuLeuLeuLeuLeu	154
Ds	1380	AGAGACCCAGTCTCCATTTGCGGTGTGACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1439
Qy	155	LeuPheIleThrSerValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGly	174
Ds	1440	TTGTTTCATCCCAAAATGCCATCTCCAGTAAGAAAGCCAGACCCAGCTCGAGGC	1499
Qy	175	AspLeuArgAsnLeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPhe	194
Ds	1500	-----TATGTTCTCTGGCGCTGAGGCGATGTGACTGACACTGACACTGTTGTGTT	1547
Qy	195	AlaGlyLysGluLeuGluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeu	214
Ds	1548	GTGAGAGAGCTCTCAGAGAAATGGGATCATGAGACTCTGACATCCCATCTACTA	1604
Qy	215	SerLysLysGluLeuProGlyValLeuGluThrGluValThrThrGlnPheIleAspGln	234
Ds	1605	GACATTGGAATGCTTATAATAATCAGGGAATCTGAAATATCTATATTTCTCCATCCA	1664
Qy	235	SerPheGlnGluPheLeuAlaLeuSerThrLeuLeuLeuLeuAspGlyGlyValPro	253
Ds	1665	TCGTTCAGGAGCTCTGTCAGCCTATCTTTATCTGCTAAAGAGCCAGCTGGACACCTCT	1724
Qy	254	ArgThrAlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisSer	272
Ds	1725	AGCCAGAGGTTTAAAGTATAGAGAACTCATGTTTGCAATTTCTTAAGAAAGTCAAAGTA	1784
Qy	273	HisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAsp	292
Ds	1785	CAGTGGATTTCTTCTGGCTCTTTCACTTTTGGCTTTTACATGAATCAGAACAAATAAAG	1844
Qy	293	IleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTrp	312
Ds	1845	CTAGAGCGGTTTCTGGCCACCAAGTGTCCCAAGAAATAAACAGTCAGTGTATCATGTC	1904
Qy	313	ValGlnGlnGlnGlyCysProGlyValAlaProGluValThrGluGlyValAlaLys	332
Ds	1905	CTGAA-----	1910

```
Db      2910 TCGAACATAAAGCTAGAGACACTGGTGAAGCTGCTGTGATGCTATATAAACATCCC 2969
QY      573 GlyCysGlyLeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuGln 592
Db      2970 AACTGCCACTTAGAGGATCTTGGGTGGGAAGCTGTGAGACTAACTGTCCTGTTGTGAG 3029
QY      593 GluLeu 594
Db      3030 GACCTT 3035
```

Search completed: October 3, 2003, 16:42:47
Job time : 3502 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OX protein - nuc-eic search, using frame_plus_p2n mode:

Run on: October 3, 2003, 11:59:35 ; Search time 6351 Seconds
'without alignments;
4025.904 Million cell updates/sec

Title:
Perfect score: 3218
Sequence:

1 MLAQPQRLLFLIGADSLPA.....ITHPALDGHFPPPKELISTF 525

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045181386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MCDELS=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029347/runat_02102003_173436_17076/app_query.fasta_1.775
-DB=GenEmbl -QPMF=fastap -SUFFIX=rge -MINMATCH=0.1 -JCOPEL=0 -JCOPEXT=0
-UNIT=54bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=humar40.cdi -LIST=45
-DOCAL=GN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=C -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US10029347@cgn_1_1_3963@runat_02102003_173436_17076 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TTIMECUT=120 -WARN_TTIMECUT=30 -THREDS=1 -XGAPEXT=0.5 -FGAPOPEXT=6
-FGAPEXT=7 -YGAPOPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

GenEmbl:

1: gb_ba:
2: gb_htg:
3: gb_in:
4: gb_on:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vi:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_ov:
22: em_pat:
23: em_ro:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:

29: em_vi:
30: em_htg_hum:
31: em_htg_inv:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rod:
36: em_htg_mam:
37: em_htg_vrt:
38: em_by:
39: em_htgo_hum:
40: em_htgo_mus:
41: em_htgo_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3218	100.0	2689	6	AX663294 Sequence
2	3218	100.0	3365	6	AX168123 Sequence
3	3207.5	99.7	2679	9	AF479748 Homo sapi
4	3207.5	99.7	2679	9	AY154461 Homo sapi
5	3170	98.5	2763	6	AX663298 Sequence
6	2882	89.6	5553	6	AX088215 Sequence
7	2882	89.6	5556	6	AX459875 Sequence
8	2882	89.6	6461	6	AX088208 Sequence
9	2622.5	81.5	95038	9	AF015416 Homo sapi
10	2619.5	81.4	143835	9	AC136475 Homo sapi
11	2617.5	81.3	158819	2	AP306283 Homo sapi
12	2209	68.6	2054	6	AX663299 Sequence
13	2020	62.8	3454	10	BC031139 Mus muscu
14	1958	60.8	2354	10	BC031519 Mus muscu
15	1746.5	54.3	238533	2	AC132968 Rattus no
16	1737	54.0	2257	10	RAT1AIR M85183 Rattus norv
17	1558.5	48.4	251206	2	AC107815 Mus muscu
18	1120	34.8	61320	2	AC135992 Homo sapi
19	835	25.9	3913	10	AF486632 Mus muscu
20	831	25.8	3221	9	AY116207 Homo sapi
21	826.5	25.7	3108	6	AX417214 Sequence
22	826.5	25.7	3300	6	AX459869 Sequence
23	826.5	25.7	3466	9	AX095460 Homo sapi
24	826.5	25.7	3563	9	AY116205 Homo sapi
25	826.5	25.7	3731	9	AY116204 Homo sapi
26	826	25.7	3395	9	AY116206 Homo sapi
27	825.5	25.7	3507	9	BC028069 Homo sapi
28	825.5	25.7	3827	9	AY095146 Homo sapi
29	825.5	25.7	3827	9	AY154467 Homo sapi
30	818.5	25.4	3237	9	AY092033 Homo sapi
31	814.5	25.3	3828	9	AF427617 Homo sapi
32	804.5	25.0	3579	9	AF468522 Homo sapi
33	800.5	24.9	3102	6	AX299762 Sequence
34	800.5	24.9	3857	6	AX299760 Sequence
35	800.5	24.9	3857	9	AF420469 Homo sapi
36	800.5	24.9	4170	9	AF410477 Homo sapi
37	762	23.7	4123	9	HS0800983 Homo sapi
38	762	23.7	4290	9	AF229059 Homo sapi
39	762	23.7	4422	6	AX089761 Sequence
40	762	23.7	4422	9	AF229060 Homo sapi
41	762	23.7	4556	6	AX089769 Sequence
42	762	23.7	5122	6	AX207661 Sequence
43	762	23.7	5306	9	BC051787 Homo sapi
44	762	23.7	5404	9	AF310105 Homo sapi
45	762	23.7	5444	9	AB023143 Homo sapi

ALIGNMENTS

RESULT 1

```

AX663294
LOCUS AX663294 2689 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 1 from Patent WO20061086.
ACCESSION AX663294
VERSION AX663294.1 GI:29163660
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Feder J., Ramanathan, C. and Mintier, G.
AUTHORS Human leucine-rich repeat containing protein, expressed
TITLE predominantly in small intestine, HLRRS1:
JOURNAL Patent: WO 02061086-A1 08-AUG-2002;
Bristo-Myers Squibb Company (US)
FEATURES
Location/Qualifiers
1..2689
/molecule="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
75..1952
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD80012.1"
/db_xref="GI:29163661"
/translation="MLAQQRLLFLDGADELPAAGPEAAPTCDPERASGARVLGG
LSKALLPTALLVITRAAAGRLGRJCSPECAEVGFSDKXKKKIFKFFRERRA
ERAFVFXENTLIFALCEVPVWIVTLRQLELGRDLSRTKTTISVYLFITSV
LSAPADGPRCLRLCARLEGRRAOFARKELELELRGSKVLTPLSKXE
LPLEFTVTVQFIDQFOEFLAALSYLEDDGVPRTAAGGVGTLIRGDAQPHSLVL
TTRELFGLLSAERMDIERHFECVXSERVKCEALRWYGGGCGFGVAPEVTEGAGL
EDTEPEEEEGEPNYPLELLYCYETQEDAFVRQALCRPEALQGVRFCEVDVAV
LSYVRCPAQALRLISLRVAAQEKKSLGRALQASUGGSGCGTQKLPASLHJ
PLFQAMTPLEHLSLTLUSHKLPDAVCRDLSAARAAPALTEGLHNLRLSEAGLX
LSEGLAWPCQVTVRQVLPQPPQVQVLMGRQSPALITLDSGQCFAPMWYLYC
AVLQHGCGLOLTSLASVELESEQSLQELQAVKRAKPLDIVITHPALDGRPPPKKEIJST
F"
BASE COUNT 534 a 861 c 844 g 450 t
ORIGIN
Alignment Scores:
Pred. No.: 9,886-236 Length: 2689
Score: 3218.00 Matches: 625
Percent Similarity: 100.00 Conservative: 0
Best Local Similarity: 100.00 Mismatches: 0
Query Match: 100.00 Indels: 0
DB: 6 Gaps: 0
US-10-029-347-2 (1-625): x AX663294 (1-2689)
QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeuProAla 20
DB 75 ATGCTGGGCGGCGCGCGCTGCTCTTCATCCGGACGCGCGCGAGCTGCGCGCG 134
QY 21 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
DB 135 CTGGGGGGCGCGCGCGCGCTGCGACAGACCCCTTCGAGCGCGCGCGCGCGCGG 194
QY 41 ValLeuGlyGlyLeuLeuSerLysLysLysLysLysLysLysLysLysLysLys 60
DB 195 GTGCTAGGCGGCGCTGTAGTAAAGCGCTGTGCGCCAGCGGCGCTCTGCTGTGAC 254
QY 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
DB 255 CCGCGCGCGCGCGCGCGCGCTGCGAGCGCGCGCTGTGTCGCCAGTGGCGCGAGCTG 314
QY 81 ArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLys 100
DB 315 CCGCGCTTCTCGACAGGACAGGACAGGAGAGAGATTTCTACAGTTCTTCGGGATCGAGG 374
QY 101 ArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120

```

```

DB 375 AGGGCGAGCGCGCTACCGCTTCGTGAAGAGAGAACGAGACGCTGTTCGCGCTGCTTC 434
QY 121 ValProPheValCysThrPileValCysThrValLeuArgGlnLeuGluLeuGlyArg 140
DB 435 GTGCCCTTCGTGTGCTGGATCGTGCCACCGTGCTGCCGACGAGCTGGAGCTCGTCGG 494
QY 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheLeuThrSerVal 160
DB 495 GACCTGTCGGCAGCGTCCAAAGACACCGCTCAGTGACTCTTTTCATCACCAGCGTT 554
QY 161 LeuSerSerAlaProValAlaAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
DB 555 CTGAGCTCGCTCCCGTAGCCGACGCGGCCCGCGTGTGCAGGGGACCTCGCGCAATCTGTGC 614
QY 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAaGlnPheAlaGluLysGluLeuGlu 200
DB 615 CGCTCGCGCGCGAGCGCTCTTCGACCGCGCGCGCGCTTCGCCGAGAGAACCTGGAG 674
QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
DB 675 CAACTCGAGCTTCGTGCTCCAAAGTGACAGCGCTGTTCCTCAGCAAAAAGAGCTGCCG 734
QY 221 GlyValLeuGluThrGluValThrTyrGlnPheLeuAspGlnSerPheGlnGluPheLeu 240
DB 735 GCGCTGCTGGACAGAGCTCCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 794
QY 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
DB 795 GCGGCACTGTCTTACCTGCTGGAGAGCGGGGTGCCAGGACCGCGCTGCGCGCTG 854
QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
DB 855 GGGACACTCTCGCTGGGCGCGCCAGCGCGCACCTGCTGCTGCTCACCACGCGCTTC 914
QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
DB 915 CTCTTCGAGCTGTGAGCGCGGAGCGGATGCGCGCACATGAGCGCGCTTCGCGCTCATG 974
QY 301 ValSerGluArgValLysGlnGluAlaLeuArgTyrValGlnGlyGlnGlyCys 320
DB 975 GTTTCAGAGCGTGTGAAGCAGGAGCGCTTCGCGGTGGGTGAGGAGCAGGAGGCTGC 1034
QY 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
DB 1035 CCGGAGTGGCACACAGAGGTGACCGAGGGGGCCAAAGGGCTCGAGGACACCGAAGGCCA 1094
QY 341 GlnGluGluGluGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
DB 1095 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1154
QY 361 GluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
DB 1155 GAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1214
QY 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
DB 1215 CAGCAGTGGCTTCCTCGCGCATGGACGTGGCTGTTCAGCTACTCTGCTGAGTGGCTGTC 1274
QY 401 ProAlaGlyGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluLys 420
DB 1275 CTTGCTGGACAGGACCTGCGGCTGTGATCAGCTGAGATTTGTTCTGCGGAGGAGAGAAG 1334
QY 421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlySerSerGlnGlyThr 440
DB 1335 AAGAGAGCTGGGAGAGCGGCTTCAGGCGAGCGCTGGGTGGGCGAGTCTCAAGGCACC 1394
QY 441 ThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGlnAlaValThrAspProLeu 460
DB 1395 ACAAACAACCTGCGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1454
QY 461 CysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAsp 480
DB 1455 TGCCATCTGAGAGCGCTCAGCTGTGCCACTGCAAACTCCCTGACGCGGTCTGCCGAGAC 1514

```

```

QY 481 LeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArg 500
Db 1515 CTTCCTGAGGCGCTGAGGCGAGCCCGCAGCTACGCGAGCTGGGCTCTCCACACAGG 1574
QY 501 LeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgVal 520
Db 1575 CTCAGTGAGCGCGGAGCTGCGTATGCTGAGTGAGGGCTAGCCCTGGCGCGAGTGCGAGGGTG 1634
QY 521 GlnThrValArgValGlnLeuProAspProGlnArgGlyLeuGlnTrpLeuValGlyMet 540
Db 1635 CAGACGCTCAGGCTACAGCTGCTGACCCCGCAGAGGGCTCCAGTACCTGGTGGGTATG 1694
QY 541 LeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaPro 560
Db 1695 CTTCGGCAGAGCCCGCCCTGACCCCTGGATCTCAGCGGCTGCCAATGCGCGCCCTC 1754
QY 561 MetValThrTrpLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSer 580
Db 1755 ATGCTGACCTACCTGTGTGCACTCTGCGAGCACAGGATGCGGCTGCAGACCCCTCACT 1814
QY 581 LeuAlaSerValGluLeuSerGluGlnSerLeuGlnGluLeuGlnAlaValLysArgAla 600
Db 1815 CTGGCTCTGTGAGCTGAGCGAGCTGCTACTACAGAGCTTCAGGCTGTGAGAGAGCA 1874
QY 601 LysProAspLeuValIleThrHisProAlaLeuAspGlyHisProGlnProProLysGlu 620
Db 1875 AAGCCGGATCTGTCATCACACACAGCGCTGAGCGGCCACCCACACCTCCCAAGGAA 1934
QY 621 LeuLeuSerThrPhe 625
Db 1935 CTCATCTCGACCTTC 1949

RESULT 2
LOCUS AX168123 3365 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 41 from Patent WO0142288.
ACCESSION AX168123
VERSION AX168123.1 GZ:1459740:
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1.
Burford,N., Baughr,M.R., Au-Young,J., Yang,J., Lu,D.A. and Reddy,E.
G-protein coupled receptors
TITLE Patent: WO 0142288-A 41 14-JUN-2001;
JOURNAL Incyte Genomics, Inc. (US)
FEATURES
     source
       1..3365
         /organism="Homo sapiens"
         /mol_type="genomic DNA"
         /db_xref="taxon:9606"
         /note="Incyte ID No: 149940831"
BASE COUNT 612 a 1106 c 1114 g 533 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1 32e-235 Length: 3365
Score: 3218.00 Matches: 625
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```

US-10-029-347-2 (1-625) x AX168123 (1-3365)

```

QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
Db 799 ATGCTGGCCAGCGCAGCGGCTGCTCTTCTATCTCGACGGCGGACAGCTGCGGCG 858
QY 21 LeuGlyGlyProGluAlaAlaProCysThrAspPheGluAlaAlaSerGlyAlaArg 40

```

```

859 CTGGGGGCCCCGAGGGCGGCGCTGCACAGACCCCTTCGAGGGCGGAGCGGCGCGG 918
QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 60
Db 919 GTGCTAGCGGGCTGCTGAGCAAGCGCTGCTGCCACGCGCTCTCTGCTGTCACACG 978
QY 61 ArgAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
Db 979 GCGCCGCGCGCGCGGAGGCTGCAGGGCGGCTGTGTTCCCGCAGTGGCGCGAGGTG 1038
QY 81 ArgGlyPheSerAspLysAspLysLysLysPheTrpLysPhePheArgAspGluArg 100
Db 1039 CGCGCTTCTCCGACAAGGACACAGAGAGTATTCTACAAGTTCTTCCGGATGAGAGG 1098
QY 101 ArgAlaGluArgAlaTrpArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
Db 1099 AGGCGCCAGCGCGCTACCGCTTCGTGAAGAGAGAACAGAGCGCTGTTCGCGCTGTGCTTC 1158
QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
Db 1159 GTGCCCTTCGTGCTGATCGTGTGCACCGTGTGCGCCAGCAGCTCGAGCTCGGTCCG 1218
QY 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
Db 1219 GACCTGTGCGCGCAGCTCCAAGACCCACCGCTCAGTGTACCTGCTTTCATCACCGAGCTT 1278
QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
Db 1279 CTGAGCTCGGCTCCGCTAGCGCGGCGCGCTGTGAGCGCGACCTGCGCAATCTGTGTC 1338
QY 181 ArgLeuAlaArgGluGlyValLeuGlyArgAlaGlnPheAlaGluLysGluLeuGlu 200
Db 1339 CGCCTGGCCCGGAGGGCGCTCTCGGACGCGAGGCGCGAGTTTCCGAGAGAACTCGAG 1398
QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
Db 1399 CAACTGGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTTTCTCAGCAAAAAGGAGCTCCG 1458
QY 221 GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
Db 1459 GCGCTGCTGGAGACAGAGGTCACCTACAGTTCTATCGACCCAGAGCTTCCAGGAGTTCTCT 1518
QY 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyValProArgThrAlaAlaGlyGlyVal 260
Db 1519 GCGGCACTGCTCTACCTGCTGGAGAGCGCGGCTGCCAGGACCGCGCTGCGCGCGTT 1578
QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
Db 1579 GGGACACTCTGCTGGGAGCGCCAGCCAGCCAGCTTGTGTGCTCACACGCGCTTC 1638
QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
Db 1639 CTCCTTCGAGCTGCTAGCGCGGAGCGGATGCGGACATCGAGCGCCACTTCGCGCTGATG 1698
QY 301 ValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlnGlyCys 320
Db 1699 GTTTCAGAGCGTGTGAAGCAGGAGGCGCTCGCGTGGGTGCAGGCGACAGGCGAGGCTGC 1758
QY 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
Db 1759 CCCGAGTGGCCAGAGGTTGACCGAGGGGGCCAAAGGGCTCGAGGACACCGAAGAGCCA 1818
QY 341 GluGluGluGluGlyGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
Db 1819 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1878
QY 361 GlnThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
Db 1879 GAGACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1938
QY 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400

```

Db 1639 CAGCGACGCGCTTCGCGCATGACGCTGCTCTTCTAGACTACTGCTGAGTGCTGC 1998
Qy ProbiaGlyGlnAlaLeuArgLeuIleSerCysArgLeuValAlaAlaGInGluLysLys 420
Db CCTCTGTCAGACGACCTGCGCTCATCAGCTGCGAGATTGGTGTGCGCAGGAGAAAG 2058
Qy LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGlyThr 440
Db AAGAAGAGCCTGGGAGCGGCTCAGCGCAGCTGGGTGGCGCAGTTCTCAAGGACCC 2118
Qy ThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeu 460
Db ACAAACAACCTGCCAGCTCCCTCTTCTCATCCACTCTTTCAGGCAATGACTGACCCACTG 2178
Qy CysHisLeuSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAsp 480
Db TGCCATCTGACGAGCTCAGCTGTGCCACTGCAAACTCTCTGAGCGGCTGCGGAGAC 2238
Qy LeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArg 500
Db CTCTCTGAGCGCTGAGGCGAGCGCCGCACTGACGAGCTGGGCTCTCTCCACACAGG 2298
Qy LeuSerGluAlaGlyLeuArgMetLeuSerGlnGlyLeuAlaIleProGlnCysArgVal 520
Db CTCACTGAGCGGAGCTGCTGATGCTGAGTGGGCTAGCTGCGCGCAGTGCAGGCTG 2358
Qy GlnThrValArgValGlnLeuProAspProGlnArgGlyLeuGlnThrLeuValGlyMet 540
Db CAGACGCTGACGGTACAGCTGCTGACCCGAGCGAGGCTCCAGTACCTGGTGGGTATG 2418
Qy LeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaPro 560
Db CTTCGCGACAGCCCGCTGACCACTGATCTCAGCGGCTGCGCACTGCCCGCCCTC 2478
Qy MetValThrThrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSer 580
Db ATGCTGACCTACCTGTGTGAGCTGCTGACGACACAGGAGTGGCGCTGCGACCTCAGT 2538
Qy LeuAlaSerValGluLeuSerGluGlnSerLeuGlnLeuGlnAlaValLysArgAla 600
Db CTGGCTCTGTGGAGCTGAGGCGAGCTGCTGACGAGCTGCTGAGGCTGTGAAAGAGCA 2598
Qy LysProAspLeuValIleThrHisProAlaLeuAspGlyHisProGlnProProLysGlu 620
Db AAGCGGATCTGGTCACTACACACCCAGCGCTGAGCGGCGCCACCTCCCAAGAGAA 2658
Qy LeuLeuSerThrPhe 625
Db CTCATCTCGACCTTC 2673

RESULT 3
AF479748
LOCUS Homo sapiens PYRIN-containing APAF1-like protein 5 mRNA, complete cds.
AF479748
VERSION AF479748.1 GI:19387135
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Primates; Carnivora; Homnidae; Homo.
REFERENCE
AUTHORS Wang, L., Yan, J., G. A., Grenier, J. M., Al-Garawi, A., Merriam, S., Lora, J. X., Geddes, B. J., Briskin, N., DiStefano, P. S., and Bertin, J. PYRAF7, a Novel PYRIN-containing APAF1-like Protein That Regulates Activation of NF-kappa B and Caspase-1-dependent Cytochrome Processing
J. Biol. Chem. 277 (33), 29874-29883 (2002)
PUBMED 2019269
REFERENCE 2 (bases 1 to 2679)
AUTHORS Bertin, J.
TITLE Direct Submission

JOURNAL Submitted (01-FEB-2002) Inflammation, Millennium Pharmaceuticals Inc., 45 Sidney Street, Cambridge, MA 02139, USA
FEATURES
source
1. .2679
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p15"
CDS
1. .2679
/note="PYRAF7; PYRIN-containing nucleotide-binding site/leucine-rich repeat family member"
/codon_start=1
/product="PYRIN-containing APAF1-like protein 5"
/protein_id="AA187105.1"
/db_xref="GI:19387135"
/translation="MDPEAPCSSTGPRFLAVARELLAALBELSDEQLKRRFHHKRDV
GPDGSIWGLERLADADLAELAQFYPPEPALEAVARKTKRADRDVAALQERRR
ORLGSGTLLSVSEYKKYREHVLQARVKNRVSVKTKRTKLLIAPESAAYE
EALGPAEPGPRARSDTHFNRLFRDEGRRLTVVLOGPAGIGKTMARKKILYD
WAAGKLYGVDFAFPMCGELLERPTGTRSLADLILDOCPDRCAPVQMLAQPRLLF
ILDGADLPALGGPEAPCTDPEAASGARVLGGLLSKALLPTALLLVITRAAGRL
ILGCSPOCAEVRGSDKKKFKYFFRDERAERAYRFRKENTFALCFVFPVW
IVTVLRQQLGRLDRLSTSTTTSTVLLFITSVLSAPVADPRQLGDLNLCFLAR
EGLVRRAGFAEKELELELRSKVQTLFSGKELPGLVEVTYQFIDQSFELAA
LSYLLEDGVPRTAAGVGTLRGDAOPHSHLVLTTRFLGGLSARMRDIERHFGCM
USERYKQALRKVQGGGCGVAPETEGAKGLDETEPEEEGESEPNVZLELLYC
LYTEQDAFARQALCFPELALQVRFCMRDVAVLSTVCRCPAGALRLSLRVLAA
GKCKKSLGKRLQASLGGSSGQGTTKQ-PASLHLPLFOAMTPDQLCHLSSLTLSHKL
PQAVRDLSEALRAAPALTEGLLHNLRLSEGLMSEGLAMPQCRVQCTVRVQLPDPQ
RGLQYLVGMRLQSPALTTDLSCCOLPAPMTYLCALVHQHCGGCTGLSLASVELSE
SLOELQAVKRAKPDVITHPALDGHQPPKELISTF"
BASE COUNT 458 a 878 c 934 g 409 t
ORIGIN
Alignment Scores:
Pred. No.: 6,21e-235 Length: 2679
Score: 3207.50 Matches: 625
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 0
Query Match: 99.67% Indels: 0
Gaps: 1
US-10-029-347-2 (1-625) X AF479748 (1-2679)
Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeuProAla 20
Db 799 ATGCTGGCCAGCGCGGCTGCTCTTCTCATCTGAGCGCGACGAGCTGCGCGG 858
Qy 2 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
Db 859 CTGGGGGGCCCCGAGCGCGCCCTGCACAGACCCCTTCGAGGGCGGCGGCGCGG 918
Qy 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
Db 919 GTGCTAGCGGGCTCTGAGCAAGCGCTGCTGCCACCGGCCCTCTGCTGCTGACCA 978
Qy 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
Db 979 CGCGCGCGCGCGCGGAGGCTGAGGCGCGCTGTGTTCGCCGAGTGCAGCGCGG 1038
Qy 81 ArgGlyPheSerAspLysAspLysLysLysTyrPheTyrLysPhePheArgAspGluArg 100
Db 1039 CGCGGCTTCTCGACAGGCAAGCAAGAAAGTATTCTTACAACTTCTTCGGGATGAGAGG 1098
Qy 101 ArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
Db 1099 AAGGCGCGCGCGCTTACCGCTTCGTAAGGAGACGAGAGCGCTGTTCGCGCTGCTC 1158
Qy 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnLeuGlnLeuGlyArg 140
Db 1159 GTGCCCTTCTGCTGGATCGTGTGCACCGCTGTGCGCGCAGCAGCTGCGTGGTGG 1218

```

141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
1219 GACCTGTGGCGCAGCTCCAGAACCCACACGCTCAGTGTACCTGCTTTTCATCACCAGCGT 1278
161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
1279 CTGAGCTCGGCTCCGGTAGCTGACGGGCCCCGGTTGCAGGGCGACCTGGCGAATCTGTGC 1338
181 ArgLeuAlaArgGluGlyValLeuGlyValArgArgAlaGlnPheAlaGlnGlyLeuGlu 200
1339 CGCTGGCCCCGGAGGGCGCTCTCGAGCAGCGGGCGCAGTTGCCGAGAGGAACCTGGAG 1398
201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
1399 CAACTGAGCTTCGTGCTCCAAAGTCACAGCCTGTTCTCAGCAAAAAGAGAGCTGCCG 1458
221 GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
1459 GGGCTGTGTGAGACAGAGGTCACCTACAGCTTCATCGACCAGAGCTTCCAGGAGTTCCTC 1518
241 AlaAlaLeuSerTyrLeuLeuGluAspGlyValProArgThrAlaAlaGlyGlyVal 260
1519 GCGCACTGTCTCACTGCTGAGGAGCGGGGTGCCAGGACCGCGCTGCGCGCGTT 1578
261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
1579 GGGACACTCTCGTGGGAGCGCCAGCCGACAGCCACTTGTGTCTCACCAGCGCTTC 1638
281 LeuPheGlyLeuLeuSerAaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
1639 CTCTTCGAGCTGTGTGCGGCGAGCGGATGCGCACATCGAGCGCCACTTCGCGTGCATG 1698
301 ValSerGluArgValLysGlnGluAlaLeuArgTyrValGlnGlnGlnGlyCys 320
1699 GTTTCAGAGCGTGTGAAGCAGGAGCGCCCTCGGTGGGTGCAGGACAGGACAGGGCTGC 1758
321 ProGlyValAlaProGluValThrGluGlyValAlaLysGlyLeuGluAspThrGluGluPro 340
1759 CCGGAGTGGCACAGAGGTGACCGAGGGGGCCAAAGGGCTCGAGGACACCGAAGAGCCA 1818
341 GluGluGlnGluGlyGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
1819 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1878
361 GluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
1879 GAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1938
381 GluArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
1939 CAGCGAGTGGCTTCTGCGCATGGACGTGGCTTCTGAGCTACTGCGTGAGGTGCTGC 1996
401 ProAlaGlyGlnAlaLeuArgLeuIleSerCysArgGluValAlaAlaGlnGluLysLys 420
1999 CTTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2058
421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGly 439
2059 AAGAGAGAGCTGGGAGAGCGCTCCAGCCAGCGCTGGTGGCGGCGAGCAGTTCCTCAGGC 2118
440 ThrThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspPro 459
2119 ACCCAAAAACACTGCCAGCGCTCCCTTCTTCATCCACTCTTCAGGCAATGACTGACCCA 2178
460 LeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArg 479
2179 CTGTGCCATCTGAGCAGCGCTCAGCTGTCCACTGCAAACTCCCTGACGCGGCTGCCGA 2238
480 AspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsn 499
2239 GACCTTCTGAGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2298
500 ArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTyrProGlnCysArg 519

```

```

1219 AGGCTCAGTGAAGCGGACTCGCTATGCTGAGTGAAGCGCTAGCTGCGCGCAGTGCAGG 2358
520 ValGlnThrValArgValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGly 539
2359 GTGACAGCGGTGAGGTACAGCTGCTGACCCCGAGGAGGCTCCAGTACCTGCTGGGT 2418
540 YetLeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAla 559
2419 ATGCTTCGGCAGAGCGCTGCGCTGACCACTGATCTCAGCGGTGCAACTGCGCGCC 2478
560 ProMetValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeu 579
2479 CCCATGGTACCTACCTGTGTGAGTCTCTGACGACCAAGGATCGGCGCTGACAGCCCTC 2538
580 SerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnGluLeuGlnAlaValLysArg 599
2539 AGTCTGGCTCTGTGGAGCTGAGCGAGCAGTCACTACAGGAGCTTCAGGCTGTGAAGAGA 2598
600 AlalysProAspLeuValIleThrHisProAlaLeuAspGlyHisProGlnProLys 619
2599 GCAAGCGGATCTGTGTATCACACCCAGCGCTGAGCGGCCACCCCAACCTCCCAAG 2658
620 GluLeuIleSerThrPhe 625
2659 GAACCTCATCTCGACCTTC 2676

RESULT 4
LOCUS AY154461 2679 bp mRNA linear PRI 20-FEB-2003
DEFINITION Homo sapiens NALP6 (NALP6) mRNA, complete cds.
ACCESSION AY154461
VERSION AY154461.1 GI:28436365
XKEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2679)
AUTHORS Tschopp, J., Martinon, F. and Burns, K.
TITLE NALP6: a novel protein family involved in inflammation
JOURNAL Nat. Rev. Cell Biol. 4 (2): 95-104 (2003)
MEDLINE 22451042
PubMed 12563287
REFERENCE 2 (bases 1 to 2679)
AUTHORS Martinon, F., Hofmann, K. and Tschopp, J.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-2002) Institute of Biochemistry, University of
Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland
FEATURES
Location/Qualifiers
1..2679
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2679
/gene="NALP6"
1..2679
/gene="NALP6"
note="member of the NALP protein family involved in
inflammation; contains NACHT, LRR and PYD domains"
/codon_start=1
/product="NALP6"
/db_xref="GI:28436366"
/translation="MDPEAPCSGPRPLAVARELLLAALBELSQEOKPRHKLRDV
GPDGRSPWGRLEPADVLAELQAQYGPPEPALEAVARKTKRADRVLAQQLRRLL
ORLQSGTTLISVSEYKDKTREVHQLHARVKNARSVKITKFTKLLTAPEAAPE
BALGPABEPGRARRDTHFNRLPRDEGRPLTVVLQGPAGIKTAAKKILVD
WAAKLVQGVDFAFNPGSELLERPGTSLADLILQCCPDGRGVPQMLAQPORLLF
ILGDRELPAVGFEAPCTDFFAASGARVRRGGLSKALLPTALLVITRAAAPGRLL
IVCTVLQQLDLRLSRTSKTSTTSTVLLFTSVSSAPVADGRPLGDLRLNCLRLAR
EVLGRRAQFAEKELELRGSKVQTLFLSKKELPGVLETVTYQFDQSFGQLAA

```


LSVLEDDGVPRTAAGVGTLLRGDAOPHSHLVLTTRFLGLLSAERMDIERHFSCM
 VSERVKDEALRWGQGGCGCPGVAPEVTEGAKGLEDETEPEEBEEBEEPNFLELLYC
 L:ETQSDAFVRQAUCRPFELALQVRFRMDVAVLSYVRCPCGAGQALRLISRLVAA
 QEKXKSLGRQLQASLGGSSQQTQOLPASLLHPLFQAMTDPCLCHSSLTLSHCKL
 PDVCRVLSLSEALRAAPALTEGLGLHNLSEAGRLMLSEGLAWPCQRTVTRVGLPPQ
 RGLQYVLMPLQSPALTTLDUSGOLPAPMVTYLCVAVLHQGGGLQTSLASVELSEC
 SLOELQAVKRAKRPDLVITHPALDGHPPPPKELISTP"

BASE COUNT: 458 a 878 c 934 g 409 t
 ORIG.N

Alignment Scores:

Pred. No.: 6,21e-235 Length: 2679
 Score: 3207.50 Matches: 625
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 0
 Query Match: 99.67% Indels: 1
 DB: 9 Gaps: 1

US-10-029-347-2 (1-625) x AV154461 (1-2679)

QY 1 Met-LeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
 DB 799 ATGCTGGCCCGACCGACCGCTGCTCTCACTCTGGACCGCGCGAGACTCCCGCG 859
 QY 2: LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAaAlaSerGlyAlaArg 40
 DB 859 CTGGGGGCCCCGAGGCCGCCCTGACACACCCCTTCGAGGGCGGCGCGCGCG 919
 QY 4: ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
 DB 919 GTGCTAGCGCGCTGCTGACCAAGCGCTGCTGCCACCGCCCTCTGCTGGTCAACAG 979
 QY 6: ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
 DB 979 CGCCCGCGCGCCCGCGGAGCGCTGACAGGCGCGCTGTGTTCCCGCACTGCGCGAGG 1019
 QY 8: ArgGlyPheSerAspLysAspLysLysLysThrPheThrLysPhePheArgAspGlyArg 100
 DB 1039 CGCGCTTCTCCGACACGACACAGAGAGATATTTCTACAGATCTTCCGGGATGAGAGG 1099
 QY 101 ArgAlaGluAaGlyAlaTyraGlyPheValLysGluAsnGlnThrLeuPheAlaLeuCysPhe 120
 DB 1099 AGGCGGAGCGCGCTACCGCTTCTGTAAGGAGMACGAGACGCTGTTCCGCGCTGTGCTTC 1159
 QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
 DB 1159 GTGCCCTTCGT 1219
 QY 141 AspLeuSerArgThrSerLysThrThrThrSerValThrLeuLeuPheIleThrSerVal 160
 DB 1219 GACCTGTCCGCGACGTCACAGACACACACGCTCAGTGTACCTGCTTTTTCACACAGGTT 1279
 QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
 DB 1279 CTGAGCTCGGCTCCGCTAGCGACCGCGCGCGCGGTTGACGGCGGCGCTCGCAATCTGTGC 1339
 QY 18: ArgLeuAaArgGluGlyValLeuGlyArgArgAlaGlnPheAaGluLysGluLeuGlu 200
 DB 1339 CCGCTGGCGCGGAGGGGCTCTCGGACGCGAGGGCGGAGTTTTCGGAAGAAGAACTGGAG 1399
 QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
 DB 1399 CAACCTGGAGCTTCGTGGCTTCCAAAGTCGAGAGCGCTGTTTCTCAGCAAAAGAGAGCTGCG 1459
 QY 221 GlyValLeuGluThrGluValThrThrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
 DB 1459 GCGCTGCTGGACACAGAGTCACTACCACTTCACTCGACACAGAGCTTCCAGGAGTTCTCTC 1519
 QY 241 AaAlaLeuSerTyrrLeuLeuGluAspGlyGlyValProAspThrAlaAaGlyGlyVal 260
 DB 1519 GCGGCACTGCTCTACTGCTGGAGACGCGCGGGTGCACAGACCGCGGCTGGCGGGCTT 1579
 QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280

DB 1579 GGACACTCTCTCGTGGGAGCCCGACCGCACAGCCACTTGGTGTCTACACAGCGCTTC 1638
 QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
 DB 1639 CTCTTCGGACTGCTGAGCGCGGAGCGGATGCGGACATCGAGCGGCACCTTCGGTGCATG 1698
 QY 301 ValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlnGlyGlnGlyCys 320
 DB 1699 GTTTCAGACGCTGTGAAGCAGGAGCCCTGCGGTGGGTGACGGGACAGGAGCGGCTGC 1758
 QY 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
 DB 1759 CCGGAGTGGCACAGAGGTGACCGAGGGGGCCAAAGGCTTCGAGGACACCGAAGAGCCA 1818
 QY 341 GluGluGluGluGluGlyGluGluProAsnTyrrProLeuGluLeuLeuTyrrCysLeuTyrr 360
 DB 1819 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1878
 QY 361 GluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
 DB 1879 GAGACGACGAGGAGACCGGTTTGCGCCCAAGCCCTGTGCCGTTCGCCGAGCTGGCGCTG 1938
 QY 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrrCysValArgCysCys 400
 DB 1939 CAGCGAGTGGCTTCTGCGCGCATGGAGCTGGTGTCTGAGGCTACTGCTGAGGCTGCTG 1998
 QY 401 ProAlaGlyGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluLysLys 420
 DB 1999 CTTGCTGGACAGGACACTGCGGCTGATCAGCTGCAGATTGGTTGCTGCGGAGGAGAGA 2056
 QY 421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyLysSerSerGlnGly 439
 DB 2059 AAGAAAGAGCTGGGAGGCGGTTCAGGCGACGCTGGGTGGCGGAGAGTCTCTCAAGGC 2118
 QY 440 ThrThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspPro 459
 DB 2119 ACCACAAAACACTGCGACGCTCCCTTCTTCATCCACTCTTCAGGCAATGACTGACCCA 2178
 QY 460 LeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArg 479
 DB 2179 CTGTGCGCTCTGAGCAGCTCAGCTGTCCCACTGCAAACTCCCTGACGCGGTCTCCCGA 2238
 QY 480 AspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsn 499
 DB 2239 GACCTTCTGAGGCGCTGAGGCGAGCCCGCCGCTGAGGAGCTGGGCTCTCTCCAAAC 2299
 QY 500 ArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArg 519
 DB 2299 AGGCTCAGTGAGCGGAGCTGCGTATGCTGAGTGAGGCGCTAGCTGGCGGAGTGCAGS 2358
 QY 520 ValGlnThrValArgValGlnLeuProAspProGlnArgGlyLeuGlnTyrrLeuValGly 539
 DB 2359 GTGCAGACGCTCAGGCTACAGCTGCTGACCCCGCCGAGCGGCTCCAGCTACCTGGTGG 2418
 QY 540 MetLeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAla 559
 DB 2419 ATGCTTCGGCAGAGCCCTGCCCTGACACCCCTGGATCTCAGCGGCTGCCCACTGCCCGCC 2478
 QY 560 ProMetValThrTyrrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeu 579
 DB 2479 CCCATGGTACCTTACCTGTGTGTCAGTCTTCGAGCACCGGATGCGGCTGCAGACCCCTC 2538
 QY 580 SerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnGluLeuGlnAlaValLysArg 599
 DB 2539 AGTCTGGCTCTGTGGAGCTGAGCGAGCAGTCACTACAGGAGCTTCAGGCTGTGAAGAGA 2598
 QY 600 AlaLysProAspLeuValIleThrHisProAlaLeuAspGlyHisProGlnProLys 619
 DB 2599 GCNAAGCCGATCTGTGTCATCACACACCGGCTGGAGCGGCCACCCCACTCTCCCAAG 2658
 QY 620 GluLeuIleSerThrPhe 625

```

Db      2659  GAACTCATCTCGACCTTC 2676
RESULT 5
AX663298
LCCUS
DEFINITION Sequence 5 from Patent WO202061086.
ACCESSION AX663298
VERSION AX663296.1 GI:29163662
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Feder, J., Ramanathan, C. and Mintier, G.
TITLE Human leucine-rich repeat containing protein, expressed
predominantly in small intestine, HLRHS11
JOURNAL Patent: WO 02061086-A 5 08-AUG-2002;
Bristol-Myers Squibb Company (US)
FEATURES
source
1..2763
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 544 a 890 c 370 g 459 t
ORIGIN
Alignment Scores:
Pred. No.: 4.67e-232 Length: 2763
Score: 3170.00 Matches: 624
Percent Similarity: 96.00% Conservative: 0
Best Local Similarity: 96.00% Mismatches: 1
Query Match: 98.51% Indels: 26
DB: 6 Gaps: 1
US-10-029-347-2 (1-625) x AX663298 (1-2763)
QY 1 MetLeuAlaGlnProGlnArGLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
DB 75 ATGCTGGCCAGCGCGAGCGACTCTCTTATCTCGGACGGCGGAGCTGCCGGCG 134
QY 21 LeuGlyGlyProGluAaAlaProCysThrAspProPheGluAaAlaSerGlyAlaArg 40
DB 135 CTGGGGGGCCCGAGGCGCGCCCTGCACACAGACCCCTTCGAGCGCGGAGCGCGCGG 194
QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
DB 195 GTGCTAGGCGGCGCTGTAGTAAGCGCTGTGCTCCACGCGCCCTCTGCTGGTGAACCG 254
QY 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
DB 255 CGCGCCGCGCGCCCGGGAGACTCAGGGCGCGCTGTGTTCCCGCAGTGGCGGAGGTG 314
QY 81 ArgGlyPheSerAspLysAspLysLysLysTyrPheTyrLysPhePheArgAspGluArg 100
DB 315 CGCGGCTCTCCGACAAAGGACCAAGAAGAATATTCTTACAAAGTCTTTCGGGGATGAGAG 374
QY 101 ArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPro 120
DB 375 AGGGCCGAGGCGCGCTACCGCTTCTGTAGAGAGAACGAGACGCTGTTCGGCGTGTCTTC 434
QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
DB 435 GTGCCCTTCGTGTGATCGTGTGCACCGTGTGCGCCAGCAGCTGGAGCTCGGTCCG 494
QY 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
DB 495 GACCTGTGGCGCAGCTCCAGACACCAACCGCTCAGTGTACCTGTCTTTTCATCAGCAGCGTT 554
QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
DB 555 CTGAGCTCGGCTCGGTAGCGGAGCGGCGCGGTGTCAGGGCGGACCTTCTGCTGTGTC 614

```

```

QY 181 ArgLeuAlaArgGluGlyValValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
DB 615 CCGCTGGCCGCGGAGGCGTCTCGGACGCGGGCGCAGTTTTCGCGAAGAACTGGAG 674
QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
DB 675 CAACTGGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTTTCTCAGCAANAAGAGAGCTGCCG 734
QY 221 GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
DB 735 GCGCTGCTGGACACAGAGTCACTACCACTACCACTACCACTACCACTACCACTACCACT 794
QY 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyVal 260
DB 795 GCGGCACTGTCTTACTCTGTGAGGACGCGGGTGCACGACCGCGGCTGCGCGGCTT 854
QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
DB 855 GGGACACTCTCTGCTGGTGGGACGCGCCAGCCGACACGCACTTGGTGTCTCACCACGCGCTC 914
QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
DB 915 CTCTTCGACTGCTCAGCGCGGAGCGGATGCGCGACATCGAGCGCACCTTCGGCTGCACTG 974
QY 301 ValSerGluArgValLysGlnGluAlaLeuArgTyrValGlnGlyGlnGlyGlnGlyCys 320
DB 975 GTTTCAGAGCGTGTGACGAGGAGGCGCTGCGGTGGTGCAGGACAGGACAGGCGCTGC 1034
QY 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
DB 1035 CCGGAGTGGCACAGAGTGCACGAGGGGCGCAAAAGGCTCGAGACACCGAAGAGCCA 1094
QY 341 GluGluGluGluGluGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
DB 1095 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1154
QY 361 GTruThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
DB 1155 GAGACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1214
QY 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
DB 1215 CAGCGAGTGGCTTCTGCGCGATGGACGTGGCTGTCTGAGCTACTGCGTGGTGGTGGTGC 1274
QY 401 ProAlaGlyGlnAlaLeuArgLeuIleSerCysArgLeuValAaAlaGlnGluLysLys 420
DB 1275 CCGTCTGGACAGGCGACTCGGCTGATCAGCTGCAGATTGGTTGGTGGCGAGAGAGAG 1334
QY 421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSetGlnGlyThr 440
DB 1335 AAGAAGAGCTGGGGAAGCGCTCCAGCGCAGCTGGGTGGCGGCGAGTTCTCAAGGCACTG 1394
QY 441 ThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeu 460
DB 1395 ACAAACAACCTGCCAGCTCCCTCTTTCATCCTCTTTCAGGCAATGACTGACCCACTG 1454
QY 461 CysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAsp 480
DB 1455 TGCCATCTGACGAGCTCAGCTGTCCCTGTCACCACTCCCTGACGCGCTTCCGCGAGAC 1514
QY 481 LeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArg 500
DB 1515 CTTCCTGAGGCGCTGAGGCGAGCGCGCAGCTGACGAGGAGTGGGCTCTCTCCACACAGG 1574
QY 501 LeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgVal 520
DB 1575 CTCAGTGGGCGGAGCTCGGTATGCTGAGTGGGCGCTAGCTGCGCGCTGCGCGAGGTG 1634
QY 521 GlnThrVal----- 523
DB 1635 CAGACGCT- CAGGTGAGCGCTGGCGTGGGAGGAGCGCTGGGATGCCCCCGCACCCACAGC 1693
QY 524 -----ArgValGlnLeuProAspProGlnArgGlyLeuGln 535

```

```
Db 1694 ACCTCTGAGGTGGCCCTCCACAGGGTACAGCTGCCTGACCCCAAGGAGGCTCCAG 1753
Qy 536 TyrLeuValGlyMetLeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCys 555
Db 1754 TACCTGGTGGTATGCTTCGGGAGAGCCCGCCCTGACACCTGGATCTCAGCGGCTGC 1813
Qy 556 GlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGly 575
Db 1814 CAACCTGCCCGCCCGCATGGTACCTGACCTGTGTGAGTCTGTGCACACACAGGAGTGGCC 1873
Qy 576 LeuGlnThrLeuSerLeuAlaSerValGlnLeuSerGlnSerLeuGlnGlnLeuGln 595
Db 1874 CTGCAGACCTCAGTCTGGCTCTGTGAGCTGACGAGCAGTCACTACAGAGGCTTCAG 1933
Qy 596 AlaValLysArgAlaLysProAspLeuValThrHisProAlaLeuAspGlyHisPro 615
Db 1934 GCTGTGAAGAGAGCAAGACCGGATCTGGTCAATCACACACCCAGCGCTGGACGGCCACCCA 1993
Qy 616 GlnProProLysGlnLeuLeuSerThrPhe 625
Db 1994 CAACCTCCCAAGAACTCATCTCGACCTTC 2023

RESULT 6
AX088215
LOCUS AX088215 5553 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 10 from Patent WO0114564.
ACCESSION AX088215
VERSION AX088215.1 GI:13397126
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Peyman,J.A., Green,C.E., Hsu,A., Browning,J.A. and Carulli,L.
TITLE Polynucleotides expressed in activated T-lymphocytes and proteins
encoded thereby
JOURNAL Patent: WO 0114564-A :0 01-MAR-2001
PUBMED 114564
FEATURES
source
location/Qualifiers
1..5553
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 512 a 1856 c 1879 g 906 t
ORIGIN
Alignment Scores:
Pred. NO.: 1.03e-209 Length: 5553
Score: 2882.00 Matches: 581
Percent Similarity: 82.82% Conservative: 7
Best Local Similarity: 81.93% Mismatches: 32
Query Match: 89.56% Indels: 90
DB: 6 Gaps: 5
US-10-029-347-2 (1-625) x AX088215 (1-5553)
Qy 1 MetLeuAlaGlnProGlnArgGlnLeuPheThrLeuAspGlyAlaAspLeuProAla 20
Db 799 ATGCTGGCCCGCCAGCCGCGCTGCTCTCATCTTGAACCGCGGAGAGCTGCCGCGC 858
Qy 21 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
Db 859 CTGGGGGGCCCCGAGCCCGCCCTTGACACAGACCCCTTCGAGGCGCGGAGCGCGCGCGG 918
Qy 41 ValLeuGlyGlyLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
Db 919 GTGCTAGCGCGGCTGCTGAGCAAGCGCTGCTGCCACCCGCTCTCTGCTGGTACCAACG 978
Qy 61 ArgAlaAlaAlaProGlyArgGlnGlnArgLeuGlnGlnCysSerProGlnCysAlaGluVal 80
Db 979 CGCGCGCGCGCGCGCGCGGAGGCTGCAGAGCGCCGCTGTGTTCGCCGAGTGGCGGAGGTG 1038
```

```
Qy 81 ArgGlyPheSerAspLysAspLysLysLysTyrPheTyrLysPhePheArgAspGluArg 100
Db 1039 CGCGGCTTCTCCGACCAAGGACCAAGAGTATTCTTCAAGTTCTTCGGGATGAGAGG 1098
Qy 101 ArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
Db 1099 AGGCGGAGCGCGCCCTTACCGCTTCGTGAAGAGAGAACGAGACGCTGTTCCGCGTGTCTTC 1158
Qy 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
Db 1159 GTGCCCTCTGTTGCTGTGATGCTGTGCACCGTGTGCGCGCAGCAGCTGCGAGCTCGTGG 1218
Qy 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
Db 1219 GACCTGTCCGCGACGCTCAAGACCAACCGTCACTGCTGCTCTTTTCATCACCAACGCTT 1278
Qy 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlnGlyAspLeuArgAsnLeuCys 180
Db 1279 CTGAGCTCGGCTCCGGTAGCCGACCGGCCCCGGTTCAGGGGAGACCTGCGCAATCTGTGC 1338
Qy 181 ArgLeuAlaArgGlyGlyValLeuGlyArgArgAlaGlnPheAlaGlnLysGluGlu 200
Db 1339 CGCCTCGCCCGGAGGCGCTCCTCGAGCGAGGCGCGCAGTTTCCCGAGAAAGAACTGGAG 1398
Qy 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
Db 1399 CAACCTGGAGCTCTGTGGCTCCAAAGTGCAGACGCTGTTCTCAGCAAAAAGAGGCTGCCG 1458
Qy 221 GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
Db 1459 GCGCTGCTGGAGACAGAGGCTCACTACCACTTCACTGACACAGAGCTTCCAGAGATTCTTC 1518
Qy 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
Db 1519 GCGGCACTGCTCTACCTGCTGGAGGACGCGCGGTGCTCCAGGACCGCGCTGCGCGCTT 1578
Qy 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
Db 1579 GGGACACTCTCTCGTGGGAGCGCCCGACCGACGACCTTGTGTCTACCAAGCGCTTC 1638
Qy 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
Db 1639 CTCTTCGGACTGCTGAGCGCGGAGGAGTGCAGGACATCGAGCGCCACTTCGCGTGCATCG 1698
Qy 301 ValSerGluArgValLysGlnGlnAlaLeuArgTrpValGlnGlyGlnGlnGlyCys 320
Db 1699 GTTTCAGAGCGGTGTGAAGCAGGAGGCGCTGCGGTGGGTGCGAGGACAGGAGCGGCTGC 1758
Qy 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
Db 1759 CCGGAGTGGCACCAGAGGAGTGCAGCGGGGGGCGCAAGGGCTCGAGGACACCAAGAGCCA 1818
Qy 341 GluGluGluGluGluGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
Db 1819 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1878
Qy 361 GluThrGlnGlnAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
Db 1879 GAGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1938
Qy 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
Db 1939 CACGAGTGGCTTCTGCGCATGAGCGTGGCTGTTCTGAGCTACTGCTGAGGTGCTGC 1998
Qy 401 ProAlaGlyGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluLysLys 420
Db 1999 CTTGCTGGACAGGACACTGCGGCTGTATGATGATGATGATGATGATGATGATGATGATG 2058
Qy 421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGlyThr 440
Db 2059 AAGAGAGGCTGGGGAAGCGGCTCCAGGCGCAGCTGGTGGTGGGCGGAGCTGGCTGGGAGCC 2118
```

```
QY 440 ----- 440
Db 2119 CAACTGGCTCCAGAGTACCCCTTCGACACCACTGCTGTGACATCTGCCCCACACCTCCA 2175
QY 440 ----- 440
Db 2179 CCAGACCTCGGCTCTCCAGGCAAGAGCTTTTGGCAGAGTTCCTTTGATATAGCTCCA 2239
QY 441 ThrLysGlnLeuProAlaSerLeuLeu----- 449
Db 2239 ATTACAGCCCTGCCAGGGGCTTGGCATCTGTTGAGAGGATGAATGCACGGTGTGGCA 2299
QY 450 -----HisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSer 464
Db 2299 GGSGCTGGGCTGGGACCCAAAGACCCATGCAATGACTGACCCACTGTGCCACTGTGAGC 2359
QY 465 SerLeuThrLeuSerHisCysLeuLeuProAspAlaValCysArgAspLeuSerGluAla 484
Db 2359 AGCCTCAGCTGTCCCACTGCCAACTCCCTGACGGCTGTGCCGAGACCTTTCTGAGGCC 2419
QY 485 LeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAla 504
Db 2419 CTGAGGGCAGCCCGCACTGACGGAGCTGGGCTCTCTCCACACACAGGCTCACTGAGGCA 2478
QY 505 GlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrValArg 524
Db 2479 GGACTGCGTATGCTGAGTGGGCTTAGCCTGGCCGACGTGCAGGGTGCAGACGGTTCAGS 2538
QY 525 ValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValClyMetLeuArgGlnSer 544
Db 2539 GTACACTGCTGACCCCGCAGGAGGCTCCAGTACCTGGTGGGTATGCTTCGGCAGAGC 2598
QY 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
Db 2599 CTTGCCCTGACCACTGGATCTCAGGGCTGCCAATCTGCCGCCCTCATGTGACCTAC 2658
QY 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAla----- 582
Db 2659 CTGTGTGCGATCTCGCAGCACCAGGATCGGGCTGCAGACCTCAGCCCTCTCGCTTCCT 2718
QY 583 -----SerValGluLeuSerGlu----- 588
Db 2719 TCTGACCGACCCCGAGTCTTCTCCGACGGTGTGAGAACCCCGGCGCGGTGGGG 2778
QY 589 -----GlnSerLeuGlnGlnLeuGlnAlaValLysArgAlaLysPro 602
Db 2779 CTGGAGTCTCGCTGGCTCGAGCGCCCGGAGCCCTCGGGCGACAGCGAGCGAGACC 2835
QY 603 AspLeuValIleThr-----HisProAla 610
Db 2839 CAGGTGGAGCGCGCGGGCGCGGGGAGGAGGAGGCGCGGAGGGAACCCCGGGC 2895
QY 611 LeuAspGlyHisProGlnProProLysGlu 620
Db 2899 CGGGGGCCCCACCCACAGCCCGCCCGGGAC 2929

RESULT 7
AX459875
LOCUS 5556 bp DNA linear PAT 28-JUL-2002
DEFINITION Sequence 18 from Patent WO0240668.
ACCESSION AX459875
VERSION AX459875.1 GI:21725646
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS 1
Tschopp,J. and Martinot,F.
TITLE Proteins and dna sequences underlying these proteins used for treating inflammations
JOURNAL Patent: WO 0240668-A 18 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
```

```
FEATURES
source
Location/Qualifiers
1..5556
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="NALP6/FI9.dna"
BASE COUNT 924 a 1856 c 1879 g 907 t
ORIGIN
Alignment Scores:
Prod. No.: 1.03e-203 Length: 5556
Score: 2982.00 Matches: 581
Percent Similarity: 82.82% Conservative: 7
Best Local Similarity: 81.83% Mismatches: 32
Query Match: 89.56% Indels: 90
Gaps: 5
US-10-029-347-2 (1-625) x AX459875 (1-5556)
QY 1 MetLeuAlaGlnProGlnArgLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
Db 799 ATGCTGGCCACGCGCAGCGGCTGCTCTTCATCTGGACGGCGGAGAGCTGCGCGG 858
QY 21 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
Db 859 CTGGGGGGCGCGGAGCGCGCCCTGCACAGACCCCTTCGAGGGCGGAGCGCGCGCGG 918
QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
Db 919 GTGCTAGCGGGCTGCTGAGCAAGGCGCTGCTGCCACGGCCCTCTGCTGTGATCACC 978
QY 61 ArgAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
Db 979 CGCGCGCGCGCGCGGGAGGCTGCAGGGCGCGCTGTGTTCCCGCAGTGCAGCGAGGTG 1038
QY 91 ArgGlyPheSerAspLysAspLysLysLysTyrPheTyrLysPhePheArgAspGluArg 100
Db 1039 CGCGGCTTCTCCGACAAAGGACAGAGAAGATATTCTACAGATTCTTCCGGGATGAGAGG 1098
QY 101 ArgAlaGluArgAlaTyrArgPheValLysGluLeuGlnThrLeuPheAlaLeuCysPhe 120
Db 1039 AGSGCCGAGCGCGCTACCGCTTCTGTAAGAGAGAACGAGAGCGCTGTTCGCGCTGTGCTTC 1158
QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
Db 1159 GTGCGCTTCTGCTGCTGGATGCTGTGCACCGTGTGCGCGCAGCAGCTGGAGCTCGGTG 1215
QY 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
Db 1219 GACCTGTGCGCACGCTCCAAAGACCACACCGTACGTGTACCTGCTTTTCATCACCAGCGTT 1278
QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAspLeuCys 180
Db 1279 CTGAGCTCGGCTCCGCTAGCGCGCGGGCGCGGTTGCGAGGCGACCTGCGCAATCTGTGC 1336
QY 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
Db 1339 CGCTTGGCCCGGAGGGGCTCTCGGACGACGAGCGCGGCGAGTTTGGCGAGAGGAAGTGGAG 1398
QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
Db 1399 CAACTGGAGTCTCGGCTCCAAAGTCAGACGCTGTTTCTCAGAAAAAGGAGCTGCCG 1458
QY 221 GlyValLeuGlnThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
Db 1459 GCGCTGCTGGAGACAGAGGTCCACTACAGTTTCATCGACGAGAGCTTTCAGGAGTTTCTC 1518
QY 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
Db 1519 GCGGCACTGTCTTACCTTGTGGAGACGGCGGGTGCACAGGACCGCGGCTGGCGGCGGT 1578
QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
Db 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
```

```
Db 1579 GGGACACTCTCGCTGGGGAAGCCAGCGGCACAGCCACTTGGTGCTCACACCGCGCTC 1638
Qy 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAsp1LeuArgHisPheGlyCysMet 300
Db 1639 CTCCTCGGACTGCTGAGCGCGAGCGGATGCGGACATCGAGCGCCACTTCGGCTGCATG 1698
Qy 301 ValSerGluAaGValLySGLuAlaLeuArgTrpValGlnGlnGlnGlnGlnGlyCys 320
Db 1699 GTTTCAGAGCGTGTGAAGCAAGAGGCCCTCGGGTGGGTGCGAGGACAGGACAGGGCTGC 1758
Qy 321 ProGlyValAlaProGluValThrGluGlyAlaLySGLyLeuGluAspThrGluGluPro 340
Db 1759 CCGGAGTGGACACAGAGGTACCGAGGGGGCCAAAGGGCTCGAGACACGAAGAGCCCA 1818
Qy 341 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 360
Db 1819 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1878
Qy 361 GluThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 380
Db 1879 GAGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1938
Qy 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
Db 1939 CAGCGAGTGCCTTCTGCGCATGCGACGTGGCTGTCTGTGAGCTACTCGGTGAGGTGCTG 1998
Qy 401 ProAlaGlyGlnAlaLeuArgIleuLeuSerCysArgIleuValAlaAlaGlnGluGlySlys 420
Db 1999 CCTCGTGACAGGACCTGCGCTGATCATGAGCTGAGATGGGTGCTGCGCAGAGAGAGAG 2058
Qy 421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGlyThr 440
Db 2059 AAGAGAGACCTGGGGAGCGGCTCAGGCCAGCTGGGTGGCGGAGCTGCTGGGAGCC 2118
Qy 440 ----- 440
Db 2119 CAACCTGGCTCCAGAGTAGTACCCTTTTCGACCACCTGCTGTGACA*CTGCCCCACACCTCCA 2178
Qy 440 ----- 440
Db 2179 CCAGACCTCGGCTCCTCCAGGCAAGGCTTTTCCAGAG*TCCTTTGAAATATAGCTCCA 2238
Qy 441 ThrLysGlnLeuProAlaSerLeuLeu----- 449
Db 2239 ATTGAGCCCCCTGCCAGCGGCTTGGCATCTGTTGAGAGGATGAATGTACAGCTTTGGCA 2298
Qy 450 -----HisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSer 464
Db 2299 GGGGCTGGGCTGGGGAGCCCAAGACCCATGCAATGACTGACCCACTGTGCCATCTGAGC 2358
Qy 465 SerLeuThrLeuSerHisCysIleLeuProAspAlaValCysArgAspLeuSerGluAla 484
Db 2359 AGCCTCAAGCTGTCTCCATCTGCAAACTCCTGACGCGGTCTCCGAGAGACCTTTCGAGG 2418
Qy 485 LeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAla 504
Db 2419 CTGAGGGGAGCCCCGCACTGAGGAGCTGGGCTCCTCCACACAGGCTCAGTGAGGCA 2478
Qy 505 GlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrValArg 524
Db 2479 GCACTGCGTATGCTGAGTGAGGGCTAGCTGCGCGCAGTGCAGGCTGCAGACGCTCAGG 2538
Qy 525 ValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGlyMetLeuArgGlnSer 544
Db 2539 GTACAGCTGCTGACCCCCAGGAGGCTCCAGTACC*GGTGGGTATGCTTCGGCAGAGC 2598
Qy 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
Db 2599 CTTGGCTGACACACCTCGATCTCAGCGGCTGCGCACTGCGCGCCCCCATGCTGACCTAC 2658
Qy 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAla----- 582
Db 2659 CTGTGTGAGTCTCTCAGCACAGGAGTGCAGGCTGCAGACCTCAGGCTCTCGCTTCTCT 2718
```

```
Qy 583 -----SerValGluLeuSerGlu----- 588
Db 2719 TCTGACCCGACCCCGAGTTCTTTCTCGGACGGTGTGCAGAACCCGCGCGCGCTGGG 2778
Qy 589 -----GlnSerLeuGlnGluLeuGlnAlaValLySArgAlaLySPro 602
Db 2779 CTGGAGTCTCGCTGCGCTCGGAGCGCCCGGAGCCCTCGGGGACAGCGAGGAGACC 2838
Qy 603 AspLeuValIleThr-----HisProAla 610
Db 2839 CAGGTGGAGGCGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2898
Qy 611 LeuAspGlyHisProGlnProProLysGlu 620
Db 2899 CGCGGCGCCCGCCACACCGCGCGCGGAC 2928
RESULT 8
AX088208 6461 bp DNA linear PAT 17-MAR-2001
LOCUS Sequence 3 from Patent WO0114564.
DEFINITION AX088208
ACCESSION AX088208
VERSION 1 GI:13397119
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Peyman,J.A., Green,C.D., Hsu,A., Browning,J.A. and Carulli,J.
TITLE Polynucleotides expressed in activated t-lymphocytes and proteins
encoded thereby
JOURNAL Patent: WO 0114564-A 3 01-MAR-2001;
Curagen Corporation (US); BIOGEN, INC. (US)
FEATURES
source
1. .6461
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1. .5556
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC34689.1"
/db_xref="GI:13397120"
/translation="MDQPEAPCSSTGPPLAVARELLLLAALSELSEQLKRFHKLKRDV
GPDKSIIPWRLERADVDLAELQAFYGPPEALEVARKTLKRDARDAVLAQLQBRK
QRLGSLTSLVSVEYKKYREHVQLHARVKERNARSVKTKRFTKLILAPESAAPE
EALGPAEPGPRARRSDTHFNLRFRDEGRRELTVLQSPAGIGTKMAKKLYD
WAAGKLYOGVDPAFMPGCGELLBRPGRSLADLIDQCPDRGADVPQMLAQORLLF
ILDGADLIPALGGPEAAPCTDPFEASGARVLGGLSKALLPTALLVTTAAAPGRL
GRLCSPOCAEVRGSDKDKKYFKYKFFDERERRAERAYFVKENETLPAICFVPVCH
IVCTVROQLELGRDLSRTSKTTTSYLLFTITSULSAPVADGPKQGLGNLRLAR
EVLGRRQAFKAELEQLERGSKVQTLFLSKKELPGVLETEVYQFIQSQFQEFLLA
LSYLDEGGVPRTAAGVGTLLRQDAQPHSLV-TTRFLGLLSAERNRDIERHFGCM
VSERVQKALRWVGQGCPCVAPVEVTEGAGLEDEEPEEGBEGBEPNYPLELLYC
LYTEODAFVROALCFPELALQVRFCRMDVAVLSYCVRCPCAGCA*RL*LCRLVAA
QKKKSLGRLQASLGGSNLQTLAPVPRPCCDICPTPPDPDLGKAPARY
PUNAPIQLPRLASVERMNVILAGAGPDKPDKTHATDECHUSSITLSHCKLPDA
VCFDLSALRAAPALTEGLLHNRLSLSEGLAWPOCRVOTVRVQDPDQRLG
QYLVGMLRQSPALTTDLGSQLPAPMYTYLCAVLHQCGGLQTLTSLSPDPTSSP
SQRCRPRGLRGLSPRSPAPSPGSDSEARTOVEAAGGAGGRRRGRPRPARGPHQ
PRPRCTAPASORPGPSRGPRTPRPLVAPRGSSFLPALAKSSTOVTPLAPCDRVG
LPLRPSMEDAGEDPTTFAHSLPSDPRLLATVYNALGTRVFDHDLHVSQVYNGA
GDDTHRALPPLNVLRLPAGMEQGLTETFDALDNTGSLFHLIEGPRFASQCIYAH
R*TL*HL*AFRVS*IA*RL*SGSP*IT*LL*RS*AF*SP*ES*PD*LD*HQ*GP*Q*AR*LY*IG*HT*
PQ*GP*Q*OE*VH*MT*P*AP*DL*TE*GE*BE*ART*DF*LV*AG*G*SA*QA*CL*TE*AL*Q*LA
RG*AL*Y*TA*QA*AO*LM*VE*GL*DV*GP*LO*LR*QA*RL*GS*LY*LL*SL*AP*Q*KA*PY*IC*H*LS
RG*GL*NG*RE*CY*MG*H*VF*WD*Q*DM*FP*SI*LM*PH*EA*AI*EY*RI*TL*DL*GA*LN*QN*
G*Y*Q*KA*F*AW*ES*AD*SL*EV*ED*Y*IG*V*Q*V*H*V*G*AV*U*LA*F*E*LY*TT*Y*TV*V*NS*RL*FA*AL
DY*V*RA*V*F*W*CS*RE*V*SW*P*RE*EK*Y*H*U*RG*V*MS*P*DE*Y*SV*NN*V*Y*TV*V*NS*RL*FA*AL
AQ*DL*G*LP*PS*Q*W*LA*V*AD*K*IV*P*F*V*E*Q*NF*H*P*E*F*D*G*P*P*RV*CF*GT*SS*RH*LV*GEV
VKQADWLLGYVPVPSLSPDVRNRKLEIYEATVSPQGPAMTWSNFAVGMMLKQADVA
```

RGLLDRSFANMAEPFKVWITENADGSGAVNFLTGMGGFLOAVFEGCTGFRVTRAGVTFD
 PVCLSGISRVSUSZIFVCGNKJNFESEDSVTVEITARAGPWAPHLEAEIWNPSQSLIS
 LLFGHKVSPFRSARIGQSPPKLPFGSSSEFPGRITFSDVRDPLQSLWVTLGSSPTE
 SLTVDPASE"

BASE COUNT 1079 a 2146 c 2150 g 1086 t
 ORIGIN

Alignment Scores:

Pred. No.: 1,25e-209 Length: 6461
 Score: 2882.00 Matches: 581
 Percent Similarity: 82.82% Conservative: 7
 Best Local Similarity: 81.83% Mismatches: 32
 Query Match: 89.56% Indels: 90
 DE: 5 Gaps: 5

US-10-029-347-2 (1-625) x AX088208 (1-6461)

Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
 Db 799 ATGCTGGCCAGCCGAGCGCTCTCTTCATCTTGACGGCGCGGACGAGCTGCCGCG 356
 Qy 21 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
 Db 859 CTGGGGGGCCCGAGGCGCGCCCTGCACAGACCCCTTCGAGGGCGCGAGCGCGCGCG 918
 Qy 41 ValLeuGlyGlyLeuLeuSerCysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
 Db 919 GTGCTAGGGGGCTGCTGAGCAGAGCGCTGCTGCCACGGCCCTCTGCTGTTGACCA 978
 Qy 61 ArgAlaAlaProGlyArgGlnGlyArgLeuLeuCysSerProGlnCysAlaGluVal 80
 Db 979 CGCCCGCGCCCGGAGAGCTCAGGGCGCGCTGTGTTCCCGCGAGTGGCGCGAGGTG 1038
 Qy 81 ArgGlyPheSerAspLysAspLysLysLysTyrPheTyrLysPheArgAspGluArg 100
 Db 1039 CGCGGGCTCTCCGACAAAGGAAGAGAAAGATTTCTACAAAGTTCTTCGCGGATGAGAG 1098
 Qy 101 ArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
 Db 1099 AGGCGGAGGCGGCGCTACCGCTTCGTCAGAGGAAACAGAGCGCTGTTCGGCGCTGTGCTTC 1158
 Qy 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
 Db 1159 GTGCGCTTCGTGTCTGGATCTGTGACCGCTGTGTCGCCAGAGCTGGAGCTCGGTGCG 1218
 Qy 141 AspLeuSerArgThrSerLysThrThrSerValTyrLeuLeuPheIleThrSerVal 160
 Db 1219 GACCTGTGCGCGAGCTCCAGACCAACCAACGCTCAGGTACCTTGCTTTTCATCACCAGCGTT 1278
 Qy 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
 Db 1279 CTGAGCTCGGCTCCGTTAGCCGACGCGCCCGGTTGCGGGCGACCTGCGGAACTGTGTC 1338
 Qy 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
 Db 1339 CGCCTGGCCCGGAGGCGCTCTCGGACGAGGGCGCGAGTTTCGCGAAGAAAGAACTGGAG 1398
 Qy 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
 Db 1399 CAACTGGAGCTTCGTGGCTCCAAAGTCAGACGCTGTCTTCAGCAAAAGGAGCTGCCG 1458
 Qy 221 GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
 Db 1459 GGCCTGTCTGGAGACAGAGGTACCTACCACTTCATCCAGCAGAGCTTCAGGAGCTTCCTC 1518
 Qy 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyValProArgThrAlaAlaGlyGlyVal 260
 Db 1519 CGGCACTGTCTACCTGCTGAGGAGCGGGGTGCCAGGACCGCGGCTGGCGCGCTT 1578
 Qy 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
 Db 1579 GGGACATCTCTGCTGGGGAGCGCCAGCCGACAGCACTTGTGCTCACACGCGCGCTTC 1638

281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
 Db 1639 CTCCTTCGAGACTCTGAGCGCGAGCGGATGCGGCACATCGAGCGCCACTTCGGCTGSCATG 1698
 Qy 301 ValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlnGlyCys 320
 Db 1699 GTTTCAGAGCGTGTGAAGCAGGAGGCCCTGCGGTGGGTGCAGGGACAGGACACGGCTGC 1758
 Qy 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
 Db 1759 CCGGAGTGGCACCAGAGGTGACCCAGGGGGCCAAAGGGCTCGAGGACACCCAGAGGCCA 1818
 Qy 341 GluGluGluGluGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
 Db 1819 GAGGAGGAGGAGGAGGAGGAGGCCCACTACCCACTGGAGTTGCTGTACTGCCCTGTAC 1878
 Qy 361 GluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
 Db 1879 GAGACGACGAGGAGGAGCGCTTTGTGCGCCAAAGCCCTGTGCCGGTTCGCCGAGCTGG 1938
 Qy 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
 Db 1939 CAGCGAGTGGCTTCTGCCGATGGACGTGGCTGTCTGAGCTACTGCGTGAGGTGCTGC 1998
 Qy 401 ProAlaGlyGlnAlaLeuArgLeuIleSerCysArgLeuValAlaAlaGlnGluLysLys 420
 Db 1999 CTTGCTGACAGGACTCGGCTGATCAGCTGCAGATTGGTTGCTGCCAGAGAGAGAG 2058
 Qy 421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGlyThr 440
 Db 2059 AAGAAAGAGCTGGGAGAGCGCTCCAGGCCAGCTGGGTGGCGGAGCTGGCTGGGGACC 2118
 Qy 440 ----- 440
 Db 2119 CAACCTGCTCCAGAGTAGTACCCTTTCGACACCCTGCTGTGACATCTGCCCCACACCTCCA 2178
 Qy 440 ----- 440
 Db 2179 CCAGACCCCTCGCTCTCCAGGGCAGGCGTTTGGCCAGAGTTCCTTTGAATATAGTCCA 2238
 Qy 441 ThrLysGlnLeuProAlaSerLeuLeu ----- 449
 Db 2239 ATTACGCCCTCGCCAGGGGCTTGGCATCTGTGAGAGGAAGAAATGTACCGGTGTGSCA 2295
 Qy 450 ----- HisProLeuPheGlnAlaValThrAspProLeuCysHisLeuSer 464
 Db 2299 GGGGCTGGGCGCTGGGGACCCAAAGACCATGCAATGACTGACCCACTGTGCATCTGAGC 2358
 Qy 465 SerLeuThrLeuSerHisCysLysLeuProAlaValCysArgAspLeuSerGluAla 484
 Db 2359 AGCCTCACGCTGTCCACTGCAAACTCCCTGACGCGGTCTGCCGAGACCTTCTGAGGCC 2418
 Qy 485 LeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAla 504
 Db 2419 CTGAGGGCAGCGCCCGCACTGACGAGCTGGGCGCTCTCTCCACACAGGCTCAGTAGGCA 2478
 Qy 505 GlyLeuArgMetLeuSerGlyGlyLeuAlaTyrProGlnCysArgValGlnThrValArg 524
 Db 2479 GGACTGCGTATGCTGAGTGGGCGCTAGCTGGCGCGAGTGCAGGGTGCAGACGGTCAAG 2538
 Qy 525 ValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGlyMetLeuArgGlnSer 544
 Db 2539 GTACAGCTGCCTGACCCCGCAGGAGGCTCCAGTACCTGTTGGGTATGCTTTCGCGCAGAGC 2598
 Qy 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
 Db 2599 CCTGCGCTGACCACTCGATCTCAGCGGCTGCCAACTGCGCGCGCCCATGGTGGTACCTAC 2658
 Qy 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAla ----- 582
 Db 2659 CTGTGTGAGTCTCTGCAGCACGAGGATGCGGCGCTGCAGACCCCTCAGCCTCTCGCTTCT 2718
 Qy 583 ----- SerValGluLeuSerGlu ----- 588

QY 435 ----- 435
Db 66346 GCAGAGATACTCTTTGTGTGAGGTGTGTGCGGGTGGAACTGTGCACCTCAGTG 66405
QY 435 ----- 435
Db 66406 TCCA^{CT}TGGCATGGGCGCAGCTCCAGACAGACACCTGGTCCCGCAGATGGGAGGCTTT 66465
QY 435 ----- 435
Db 66466 GACCACCTCTTTGCAAGGCAACCACTGAATTCCTATCTTTTCCCGCAGGTGTCTCACC 66525
QY 435 ----- 435
Db 66526 ATGCCACGGTCAGCCACCCACCACTATGAAACTTCCTAGTAGTGGCCCTACAGCTCTG 66585
QY 435 ----- 435
Db 66586 GCGCAGCCCTTCTCCCATTTTACTTTTCTGACAA^{TA}AGCTCTGGAACACACGATGCCCTC 66645
QY 435 ----- 435
Db 66646 TGTCCAGACCTCTGACCTTGCCCTTCCTCCCTCACC^{TA}AA^{TA}TTCTGTTTGTCTGCTGTC 66705
QY 435 ----- 435
Db 66706 CTGCAAGTCTGCCCTGGCTCCCCAGCTGGCTGGGACCCCACTGGCTCCAGAGTACCC 66765
QY 435 ----- 435
Db 66766 T^{TT}TCGACCACTCTGTGT^{GACATCT}GCCCCACACCTCCACAGACCTCGGCTCTCTCCAG 66825
QY 435 ----- 435
Db 66826 GCGAAGGCTTTGCCAGAGTTCCTTTGAATATAGCTCCAA^{TT}CGACCCCTGCCAGGGCG 66885
QY 435 ----- 435
Db 66886 TTGGCATCTGTTGAGAGGN^{TA}ATGTACGGTGT^{TT}GGCAGGGCTGGSCCTGGGACCCCA 66945
QY 435 ----- 435
Db 66946 AAGACCCATGTAGGGGA^{GT}GGGTGATACCGTCCCACTGACCTCACTCAGACACCCCACT 67005
QY 435 ----- 435
Db 67006 GGGATGATGCCCTGGGCCACAGATGAGTGGTTTGTCTGGG^{TT}GGTTGGGGAGTGTGA 67065
QY 435 ----- 435
Db 67066 AGGGGTGGCTCAGCAAGGAGGTGAGGAGGGGCGAGCTCTGAGACCCAGGACTACAG 67125
QY 435 ----- 435
Db 67126 ATAGACAGAGTTCCTTTCCGGCAGGCACCTGTGAGGAGCAGGTGGGCTTCACCT^{TT}CCCTCT 67185
QY 436 ----- 452
SerSerGlnGlyThrThryslglnLeuProIlaSerLeuLeuHisProLe
|||||
67186 CTGTCCCAGCAGTCTCTCAAGGCACCAAAACAACTGCCAGCCTCCCTTCTCTCA^{TT}CCACT 67245
QY 452 uPheGlnAlaMetThrAspProLeuCySHisLeuSerSerLeuThr----- 467
|||||
67246 CTTTCAGGCAATGACTGACCCCACTGTGCCATCTGAGCAGCCTCAGCTGNG^{TT}GGCCACACC 67305
QY 467 ----- 467
Db 67306 CCCAGCTCTTCCCAGGAGCGGGCTGAGAGCAGGATGCC^{TT}GGGACAGACGGAAGTATG 67365
QY 467 ----- 467
Db 67366 ACCTAGAAAGAGGCTGAGGGTCCAGACAGAGCCTGAGSCCTGTGGGTCTCGSAGATG 67425
QY 467 ----- 467

Db 67426 GCAGGCTCTGCCACACAGACAGAGCAGGGAGGGGATGAGGCTCTGGGGCTTCATTACC 67485
QY 467 ----- 467
Db 67486 AGGTTAGTACTGTAGAAAGGTTGGTCTCTGGTGTGAAAGCAGCTAAACACACTG 67545
QY 467 ----- 467
Db 67546 AGGAGGGATTAAGAAGACAAAGAGTGTATTTTCTTCTCAAGSGAGGCTGCCGTTGGAA 67605
QY 467 ----- 467
Db 67606 ACTGCTGGGTTCTCAGACAGTGGAGGAAGGTGT^{TT}CAAGCGCAGGCTCCAAGGTCAGGC 67665
QY 467 ----- 467
Db 67666 TGTCTGCAAAATGCTAGCCGGCTTCTGGTGCCCACTAGCAAGTCCCA^{TT}CTGTCCGAGCTTC 67725
QY 467 ----- 467
Db 67726 CATTTCTCCCTGTGAAAT^{GG}AGGCCAGTAGATCCAGAGAGTTGGGCTGAGATTTAATGT 67785
QY 467 ----- 467
Db 67786 GGCTACACATGTACAGTCTCTTTTCTTTTCTTTTCTTTT^{TT}TTT^{TT}GAGACGGAGCCTCGCTCTGTC 67845
QY 467 ----- 467
Db 67846 GCGCAGGCTGGAGTGCAGTGGTGTGATCTCAGCTACT^{TT}CAAGCTCCGCTCCCGGTTTC 67905
QY 467 ----- 467
Db 67906 ACGCCATTTCTCTGCCCTCAGCCTACCGAGTAGCTGSGA^{TT}CTACAGGTGCCGCCCACTACGC 67965
QY 467 ----- 467
Db 67966 CCGGCTAATTGTTGTATTTTGTAGTAGACAGGG^{TT}TTCACCGGTTAGCCAGGATAGTC 68025
QY 467 ----- 467
Db 68026 TCGATCTCTGACCTCGTGATCCACCCGCTCTGCCCTCCAAAGTGTCTGGATCACAGGC 68085
QY 467 ----- 467
Db 68086 ATGAGCCACCGCGCCCGCCACATGTACAGTTC^{TT}TACATGGCAAT^{AA}CTT^{TT}GCCAGTCACT 68145
QY 467 ----- 467
Db 68146 CTTAGCTACCCCAAGTGAGCAGAAAGAGAGAAATCTAGGA^{CT}AGTACTGATGGTG 68205
QY 467 ----- 467
Db 68206 AAACAGTCCATCTCTAGTTGATGGCAACTCGGCAAA^{TT}AGCTTTGATATGGGGTGGGG 68265
QY 467 ----- 467
Db 68266 ATT^{TT}GGCAATCTGGAGATAAACA^{TA}AGAACCAAAAGGTGATCACATTTGGCCCCAGCAAT 68325
QY 467 ----- 467
Db 68326 TCTGCTGGTAAGTAAGATGTTCCAGCACCACTATT^{TT}TATATGGAA^{AAAA}AAAAAGGA 68385
QY 467 ----- 467
Db 68386 GAAACAACAACACTACCCAACCAT^{TT}AGGGAATAGCTGACTAAAT^{AA}CTGAGTTGSCCAC 68445
QY 467 ----- 467
Db 68446 CTCACTTTGGTGGACGAT^{TT}GTGGGCAAAACATGTTGAAACACAGTCTCAGCTCTCCGAAGA 68505
QY 467 ----- 467

Db 68506 GCTCGTAGAGTGTCCACGSGAGAGTGAACACAGCTCCCCGGGGTGCATGGTGACAA 68565
 Qy 467 ----- 467
 Db 68566 AGGCAGGCGAGGCTTGGCGGGTCCAGGGAAGGATGTGGCAACACATCTCTCAGCT 68625
 Qy 467 ----- 467
 Db 68626 GAACCTTGGGCGCCAGCGGCGAGCGGGCATTTTGTGCGAGTTGGAGTGGTGGTGGTGG 68695
 Qy 468 ----- LeuSerHisCysLysLeuProA 475
 Db 68686 CGCGCTCAGGTAAATCTCTGTGCTTCTTGACCACAGCTGTGCCACTGCMAACTCCCTG 68745
 Qy 475 spAlaValCysArgAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuS 495
 Db 68746 ACGCGGTCTCCGAGACCTTCTGAGCGCTGAGGCGAGCCCGCCGCACTGAGGAGCTGG 68805
 Qy 495 lylLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAla 515
 Db 68806 CCCTCTCCACACAGAGCTCAGTGAGGACGAGTGGGTATGCTGAGTGAGGCGCTAGCT 68865
 Qy 515 rpProGlnCysArgValGlnThrVal ----- 523
 Db 68866 GCGCGAGTGCAGGTCACAGACGCT-CAGGTGAGGCTGGCTGGGAGGACCGTGGAT 68924
 Qy 524 ----- ArgValGlnLeuProAsp 530
 Db 68925 GCCCCGCCACCCAGAGCTCTAGTTCAGTTCAGGTCGGCCCTCCACAGGGTACAGTGCCTGACC 68984
 Qy 530 roGlnArgGlyLeuGlnThrLeuValGlyMetLeuArgGlnSerProAlaLeuThrThL 550
 Db 68985 CCCAGCAGGGCTCCAGTACTGTGTGGTATGCTTGGCAGAGCCCTGCCCTGACCAACC 69044
 Qy 550 euAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuS 570
 Db 69045 TGGATCTCAGCGGTGGCAATGCCCCGCCATGCTGGTACCTTACCTGTGTGAGTCTGCTG 69104
 Qy 570 lnhAsGlnGlyCysGlyLeuGlnThr ----- 578
 Db 69105 AGCACAGGATCGGGCTCGAGCCCTCAGGTGGAGGAGGGTGGGAAGGGTGTGGG 69164
 Qy 578 ----- 578
 Db 69165 GACACAGCTGTCAACCGCGGAGGGGGTGGCTGGGGGCTTCCCTGGACCCCTGAA 69224
 Qy 578 ----- 578
 Db 69225 GAGGGGCCCTCCACAGACCTTCCATGTGACTGAGCTCAACACACTGACCTGGAGCCCA 69284
 Qy 578 ----- 578
 Db 69285 GCCCTGACAGCCCTGCCCTCTGAAGCCCTGGCCACAGCTCGGGGCTTGACCACTTGTC 69344
 Qy 578 ----- 578
 Db 69345 TGACCAAAATTTATCTGTGCTGTGTCGATGGCAGCACCTGGGCGCAGTGGCCCGA 69404
 Qy 578 ----- 578
 Db 69405 TACCCTTTGCCACCCCAAGGTCGGGCTGTGCCCCCTGCCAGGCTCTGACAAACCT 69464
 Qy 578 ----- 578
 Db 69465 CTCAACTGACACACACTGTCTCTGAGTGTCTGTGGACAGCAGCAGGTGACAGCCAGTCA 69524
 Qy 578 ----- 578
 Db 69525 CTGCCCGCAGCCCGCCGCCCCACAGGCACTGCCGCTCACTGCCCCGCGCGCTCCCC 69584
 Qy 578 ----- 578
 Db 69585 CCACGGCACTGCCCCCAAGCCCTGGTGTCTTCCCCCAGCGCTGACCCCGCTTCTGCTCT 69644

Qy 579 ----- Leu-SerLeuAlaSerValGluLeuSerClnSerLeuGlnGluLeuGln 595
 Db 69645 GCGTGTGGCTTGGAGTCTGTGGAGCTGAGCAGCAGTCACTACAGGAGCTTCAG 69704
 Qy 596 AlaValLysArgAlaLysProAspLeuValIleThrHisProAlaLeuAspGlyHisPro 615
 Db 69705 GCTGTGAGAGAGCAAGCCGGATCTGGTATCACACACCCAGCGCTGGAGCGCCACCA 69764
 Qy 616 GlnProProLysGluLeuIleSerThrPhe 625
 Db 69765 CAACCTCCCAAGGAAGTCACTCGACCTTC 69794

RESULT 1C

AC136475 143835 bp DNA linear PRI 30-MAY-2003
 Homo sapiens chromosome 11, clone RP11-326C3, complete sequence.
 AC136475
 AC136475.6_GI:31193978
 HTG.
 Homo sapiens (human)
 ORGANISM

REFERENCE

1 (bases 1 to 143835)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Birren, B., Nusbaum, C. and Lander, E.

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 143835)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kanat, A., Karatas, A., Kellis, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., MacLean, C., Macdora, P., Major, J.,
 Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T.,
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Submitted (02-NOV-2002) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 143835)

Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N.,
 Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cooke, P., Cooke, P., Corum, B., Dearellano, K.,
 Diaz, J. S., Dodge, S., Socie, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kanat, A., Karatas, A., Kellis, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ranasingh, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
 Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome

REFERENCE
AUTHORS

Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 143835)
 Arrhen, B., Nussbaum, C., Lander, E., Abouelheil, A., Allen, N.,
 Anderson, S., Arachchi, H., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Boukhgaiter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collamore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, J., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ranasamy, J., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomant, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Testave, S., Theodore, J., Tephani, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (30-MAY-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 30, 2003 this sequence version replaced gi:29469599.
 All repeats were identified using RepeatMasker:
 Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submission@genome.wi.mit.edu
 ----- Project Information
 Center project name: L28535
 Center clone name: 325_C_3

Only the first 43.8 kilobases of this clone are being submitted.
 The remainder overlaps accession number AC138230 [WICGR project
 L29013].

FEATURES

source

1..143835
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="Genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-326C3"
 /clone_lib="RP11 Human Male BAC"

repeat_region 1712..2006
 /rpt_family="AluSg"
 repeat_region 2347..2703
 /rpt_family="THE1B"
 repeat_region 4157..4302
 /rpt_family="L1ME4a"
 repeat_region 4435..4455
 /rpt_family="AT_rich"
 repeat_region 4528..4591
 /rpt_family="AT_rich"
 repeat_region 4592..4901
 /rpt_family="AluSx"
 repeat_region 4902..4968
 /rpt_family="AT_rich"
 repeat_region 4979..5091
 /rpt_family="L1ME4a"
 repeat_region 5154..5251
 /rpt_family="L1ME4a"
 repeat_region 5620..5683
 /rpt_family="L1ME4a"
 repeat_region complement(5684..5954)
 /rpt_family="AluJo"

repeat_region 5955..6050
 /rpt_family="L1ME4a"
 repeat_region complement(6059..6206)
 /rpt_family="MER46C"
 repeat_region 6349..6578
 /rpt_family="L1ME4a"
 repeat_region 6792..7203
 /rpt_family="L2"
 repeat_region complement(8884..9052)
 /rpt_family="L1MC5"
 repeat_region 9225..9287
 /rpt_family="MER68-int"
 repeat_region 9676..9986
 /rpt_family="AluSg"
 repeat_region 9994..10032
 /rpt_family="AT_rich"
 repeat_region 10182..10417
 /rpt_family="L1MED"
 repeat_region complement(10418..10946)
 /rpt_family="L1TR59"
 repeat_region complement(10947..11251)
 /rpt_family="AluSx"
 repeat_region complement(11252..11272)
 /rpt_family="L1TR59"
 repeat_region complement(11290..11595)
 /rpt_family="AluSp"
 repeat_region 11702..12014
 /rpt_family="AluSg"
 repeat_region 12047..12137
 /rpt_family="L1M4"
 repeat_region complement(12159..12465)
 /rpt_family="AluSx"
 repeat_region 12465..12605
 /rpt_family="L2"
 repeat_region 12763..12832
 /rpt_family="L2"
 repeat_region 12898..13715
 /rpt_family="L1MB5"
 repeat_region 13715..14007
 /rpt_family="AluSg"
 repeat_region 14018..14320
 /rpt_family="AluSx"
 repeat_region 14321..14603
 /rpt_family="L1MB5"
 repeat_region 14611..14772
 /rpt_family="AluSg/x"
 repeat_region 15386..15685
 /rpt_family="AluSx"
 repeat_region 15701..15984
 /rpt_family="AluJo"
 repeat_region complement(16654..16949)
 /rpt_family="AluSx"
 repeat_region 17027..17095
 /rpt_family="L1MC5"
 repeat_region complement(18948..19258)
 /rpt_family="AluY"
 repeat_region 19299..19409
 /rpt_family="L2"
 repeat_region complement(19426..19719)
 /rpt_family="AluSg"
 repeat_region complement(19742..20040)
 /rpt_family="AluJo"
 repeat_region 20207..20435
 /rpt_family="L1MC/D"
 repeat_region 20906..21032
 /rpt_family="(CCCCCG)n"
 repeat_region complement(21181..21341)
 /rpt_family="MIR"
 repeat_region complement(22663..22828)

Alignment Scores:
 Pred. No.:
 Score:

6.83e-188 Length:
 2619.50 Matches:
 143835
 623

	Percent Similarity:	39.20%	Conservative:	:	
	Best Local Similarity:	39.13%	Mismatches:	1	
	Query Match:	81.40%	Indels:	968	
DB:		9	Gaps:	4	
JS-1C--029-347-2 (-625) x AC136475 (1-143835)					
QY	1	MetLeuAlaGlnProGluArgLeuLeuPheIleAspGlyAlaAspGluLeuProAla	20		
DB	85077	ATCGTGGCCACAGCCGAGCGGTCTTCATCTTGAGACGGCGGACAGACTGCCGGCG	65136		
QY	21	LeuGlyGlyProGluAlaProCysThrAspProPheGluAlaAlaSerGlyAlaAsq	40		
DB	65137	CTGGGGGGCCCCAGGCCTGCAACACCCCCTGACACACCCCTTGAGGGCGGACGGCGCGGG	65196		
QY	41	ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr	60		
DB	65197	GTGCTAGGGGGCTCTCGAGCAAAGCGCTGCTGCCACAGCCCTCTCTGCTGGTGAACACG	65256		
QY	61	ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal	80		
DB	65257	CACGCGCGCCCGCCGGAGAGCTGCGAGGGCGCCCTGTGTCCCGCAGTGGCGGAGGTG	65316		
QY	81	ArgGlyPheSerAspLysAspLysLysTyrsPheTyrLysPhePheArgAspGluArg	120		
DB	65317	CSGGGCTTCCGACACAGACAAGAAGAAATTTCTACAGATAATTTCCGGGATGAGAGG	65376		
QY	101	ArgAlaGluArgAlaTyrArgPheValLysGluLysGlnThrLeuPheAlaLeuCysPhe	120		
DB	65377	AGGGCGAGCGCGCTTACCCTTCTGTAAGGAGAACAGAGACGCTGTTCGGCGCTGTGCTTC	65436		
QY	121	ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg	140		
DB	65437	STGCCCTTCTGTGTGGTATGTGTGACACGTGCTGCGCCAGCAGCTGGAGCTCGTCTGG	65496		
QY	141	AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal	160		
DB	65497	GACCTTCGCGACGCTCCAAGACCAACACGCTCAGTGTACCTGTCTTTCATCACACGCTT	65556		
QY	161	LeuSerSerAlaProValAlaAspGlyPrcArgLeuGlnGlyAspLeuArgAsnLeuCys	180		
DB	65557	CTGAGCTCGCTCCGGTAGCCGACGGGCCCGCTTGTCAGGCGACCTCGCGCATCTGTGC	65616		
QY	181	ArgLeuAlaArgGluValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu	200		
DB	65617	CGGCTGCCCGGAGGGCGTCTCGAGCGCAGGGCGCAGTTTGCAGAAAGAGAGCTGCCG	65676		
QY	201	GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro	220		
DB	65677	CAACTGGAGCTCTGTGGCTCCAAAGTGTGACAGCTGTGTCTCAGCAAAAAGAGAGCTGCCG	65736		
QY	221	GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu	240		
DB	65737	GGCGTCTCGAGACAGAGTCACTTACAGTTCATCGACACAGAGCTTCCAGGAGTTCCTC	65796		
QY	241	AlaAlaLeuSerTyrLeuLeuGlnAspGlyGlyValProArgThrAlaAlaGlyVal	260		
DB	65797	CGGGCACTGTCTTACCTGCTGAGAGACGGCGGGTGGCCACAGACCGCGCTTGGCGGCGTT	65856		
QY	261	GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe	280		
DB	65857	GGGACACTCTGGTGGGACCCCGCAGCGCACACCACTTGGTGTCTCACACGGCTTC	65916		
QY	281	LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet	300		
DB	65917	CTCTCGGACTGTGAGCGCGGAGGATGCGCGACATCGAGCGCCACTTCGGCTGCATG	65976		
QY	301	ValSerGluArgValLysGlnAlaLeuArgTrpValGlnGlyGlnGlyGlnGlyCys	320		
DB	65977	GTTTTACAGCGTGTGAAGCAGAGGCGCCCTGCGGTGGTGTGGTGGTGGGAGGTGTGA	66036		
QY	321	ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro	340		

QY	435	-----	435	QY	467	-----	467
Db	67177	ATAGACAGAGTCTTTTCGGCAGGCACTGTGAGGAGCAGGTGGGCTTCACCTTCCCTCT	67236	Db	68257	AAACAGGTCCATCTCCTAGTTGATGGCACTCGCAAAATTAGCTTGAATATGGGGTGGGG	68316
QY	436	-----SerSerGlnMlyThrThrLysGlnLeuProAlaSerLeuLeuHisProLe	452	QY	467	-----	467
Db	67237	CTCTCCAGCAGTCTCTCAAGGCACCAAAACAACCTGCCAGGCTCCCTTCTTCATCCACT	67296	Db	68317	ATTTGGCAATCTGGAGATAACATAAGAACCAAAAGGTGATCACATTTGGCCCAAT	68376
QY	452	uPheGlnAlaMetThrAspProLeuCysHisLeuSerSerLeuThr	467	QY	467	-----	467
Db	67297	CTTTAGGCAATGACTGACCTCACTGTGCCATCTGAGCAGCCTCAGGTGAGTGGCCACACC	67356	Db	68377	TCTGTGTAAGTAAGATGTTTCAGCACACCATTTATTTATAATGGAAAAAAGGA	68436
QY	467	-----	467	QY	467	-----	467
Db	67357	CCCAGCTCTTCCACGGAGGGGCTGAGAGCAGGATGCCCTGGGCAGAGACGSAAGTATG	67416	Db	68437	GAACAACACAACACTACCCAAACCATTAGGGAATAGCGACTAAATAAATACTGAGTTGGCCAC	68496
QY	467	-----	467	QY	467	-----	467
Db	67417	ACCTAGAAAGAGGCTGAGGGGTACAGACAGACCTGAGGCTGTGGTGTGGAGATG	67476	Db	68497	CTCATCTTGGTGGACGATTGTGGGCAACATGTTGAAACACGCTCTCAGCTCTCCGAAGA	68556
QY	467	-----	467	QY	467	-----	467
Db	67477	GCAGGCTCTGCCACACAGCAGCAGGAGGAGGGGATCAGGGTCTGGGGCTTCATTACC	67536	Db	68557	GCTGTAGAGAGTGTGCGAGGAGAAATGAACACAGCTCCCGGGGTGCATGGTGACAA	68616
QY	467	-----	467	QY	467	-----	467
Db	67537	AGTTAGGTACTGGTAGGAAAGGTTGGTCTTGGTGGTGGAAAGCAGCTAAGCACACTG	67596	Db	68617	AGGCAGGGCAGGCTCTGGCAGGTCCAGGGAAGATGTGSCACCAACACATCTCTCAGCT	68676
QY	467	-----	467	QY	467	-----	467
Db	67597	AGGGAGGGATTAAAGAAGACAAAGTGTATTTTCTCTCTGAAGGAGGCTGCCGTGGAA	67656	Db	68677	GAACCTGGGCCACCGGCGAGCGGCATTTTGTGCACTTGGAGGTGTGTGGTGGCTGAG	68736
QY	467	-----	467	QY	468	-----LeuSerHisCysLysLeuProA	475
Db	67657	ACTGCTGGGTTCTCAGAAGCAGTGGAGGAAGTGTTCAGAGCAGGCTCAAAGTCAGGC	67716	Db	68737	GCGCTCAGGTAATACTCTGTGCTTCTTGACCACAGGCTGTCCCACTGCAAACTCCCTG	68796
QY	467	-----	467	QY	475	spAlaValCysArgAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuG	495
Db	67717	TGCTGCAAAATGTAAGCAGTCTTCTGCGTGGCCAGTAGCAATCCATCTGTCCGAGCTTC	67776	Db	68797	ACGGGTCTCCGAGACCTTCTTGAGGCTTGAGGCGAGCCCCCGCCTGACGGAGCTGG	68856
QY	467	-----	467	QY	495	lyLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaT	515
Db	67777	CACTTCTCCCTCCCTGAAATGGAGCCAGTAGATCCAGAGAGTGGGGTGAGATTTAATGT	67836	Db	68857	GCTCTCTCCACACAGGCTCAGTGAGCGGAGCTGCGTATGCTGAGTGAGGGCTAGCCCT	68916
QY	467	-----	467	QY	515	rpProGlnCysArgValGlnThrVal	523
Db	67837	GGCTACACATGTACAGTCTTTTTTTTTTTTTTTTTTTTGTGAGCGGAGCTCGCTCTGTC	67896	Db	68917	GGCCGAGTGCAGGGTGCACACGGT-CAGGTGAGGCTGSCCTGGGAGGACCGTGGAT	68976
QY	467	-----	467	QY	524	-----ArgValGlnLeuProAspP	530
Db	67897	GCCGAGGCTGGAGTGAGTGTGTGATCTCAGCTCACTGCAAGCTCCGCTCCCGGGTTC	67956	Db	68976	GCCCCCGCCACCCAGCAGGCTCTGAGTCCGCTCCACAGGGTACAGCTGCTGACC	69035
QY	467	-----	467	QY	530	roGlnArgGlyLeuGlnTyrLeuValGlyMetLeuArgGlnSerProAlaLeuThrL	550
Db	67957	ACGCCATTCTCCTGCTCAGCCTACTGAGTAGCTGGGACTACAGGTGCCCGCCACTACGC	68016	Db	69036	CCACAGCAGGCTCCAGTACCTGTGGTATGCTTGGCAGAGCCCTGCCCTGACCAACC	69095
QY	467	-----	467	QY	550	euAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuG	570
Db	68017	CCGGCTAATGTTGTGATTTTATAGTAGACACAGGGTTTACCGGGTTAGCCAGGATGTC	68076	Db	69096	TGGATCTCAGCGGCTGCCAACTGCCCGCCCCCATGTGTGACCTACCTGTGTGTCATCTGC	69155
QY	467	-----	467	QY	570	lnHisGlnGlyCysGlyLeuGlnThr	578
Db	68077	TGGATCTCCTGACCTCGTGTATCCACCGGCTCTGGCTTCCCAAAAGTGTGGGATCAGGC	68136	Db	69156	AGNACAGGGATGGGGCTTCACACCTCAGTGGAGGAGGGGTGGGAAGGGTGTCTGGG	69215
QY	467	-----	467	QY	578	-----	578
Db	68137	ATGAGCCACCGGCGGCAACATGTACAGTCTTTACATGGCAATAACTTGCAGTCACT	68196	Db	69216	GACACAGCTGTCAACCCCGGAGGGAGGGTGCCTGGGGGCTCCCTCGGACCTGAA	69275
QY	467	-----	467	QY	578	-----	578
Db	68197	CTTAGCTACCACTGAGCAGAAAGGAAGGAATCTAGGAACCTAGTACTGATGTG	68256	Db	69276	GAGGGGCTCCCGAGAGACCTCCCATGTGACTGTGACTCAACACACTGACCTGGGAGCCCA	69335
				QY	578	-----	578

* 154136 154235: cap of 100 bp
* 154236 155662: contig of 1427 bp in length
* 155663 155762: cap of 100 bp
* 155763 156819: contig of 1057 bp in length.

FEATURES

Location/Qualifiers
1..156819
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p"
/clone="RP11-326C3"
1..21001
/note="assembly_fragment"
21102..40714
/note="assembly_fragment"
40815..58430
/note="assembly_fragment"
58531..66754
/note="assembly_fragment clone_end:SP6 vector_side:left"
66855..81334
/note="assembly_fragment"
81235..89761
/note="assembly_fragment"
89864..101889
/note="assembly_fragment"
101990..112625
/note="assembly_fragment"
112726..119802
/note="assembly_fragment"
119903..125181
/note="assembly_fragment"
125282..129184
/note="assembly_fragment clone_end:T7 vector_side:right"
129285..137260
/note="assembly_fragment"
137361..143542
/note="assembly_fragment"
143643..146891
/note="assembly_fragment"
146992..149568
/note="assembly_fragment"
149669..152678
/note="assembly_fragment"
152779..154135
/note="assembly_fragment"
154236..155662
/note="assembly_fragment"
155763..156819
/note="assembly_fragment"
37089 a 39887 c 40186 g 37757 t 1800 others

Alignment Scores:

Pred. No.: 1,08e-187 Length: 156819
Score: 2617.50 Matches: 622
Percent Similarity: 33.13% Conservative: 1
Best Local Similarity: 35.07% Mismatches: 2
Query Match: 81.34% Indels: 969
DB: 2 Gaps: 4

US-10-029-347-2 (1-625) x AP006283 (1-156819)

QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuLeuAspGlyAlaAspGlnLeuProAla 20
DB 12647 ATGCTGGCCAGCCGAGCGGTGCTCTTCATCTGGACCGCGCGAGCGAGTGGCGCG 12766
QY 21 LeuGlyGlyProGluAlaAProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
DB 12707 CTGGGGGGCCCCAGGCGCCGCCCTTCGACAGACCCCTTCGAGCGCGGAGCGCGCGG 12766
QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60

DB 12767 GTCTAGCGGGCTGCTGAGCAAGCGCTGCTGCCACGGCCCTCTGCTGGTGACCACG 12826
QY 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
DB 12827 CGCGCGCGCGCCCGGAGGCTGCGAGGCGCGCTGTTCCCGCAGTGCSCGAGGTG 12886
QY 81 ArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysPheArgAspGluArg 100
DB 12887 CGCGGCTTCTCCACAGGACAAAGAAAGATATTTCTACAAGTATTTCCGGATGAGAGG 12946
QY 101 ArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
DB 12947 AGGCGGAGCGCGCTACCGCTTCGTGAAGGAGAACGAGCGCTGTTCCGCTGTCTTC 13006
QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
DB 13007 GTGCCCTTCGTGTGCTGATCGTGTGCACCGTCTCGCAGCAGTGAGCTCGTCCG 13066
QY 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
DB 13067 GACCTGTCCGCCACGTCCAGACCACACGTCAGTGTACTGCTTTTTCATCACACGCTT 13126
QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
DB 13127 CTGAGCTCGGCTCCGCTAGCCGACCGCGCGCTTGCAGGGCGACCTGCGCAATCTGTGC 13186
QY 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
DB 13187 CGCCTGGCCCGCGAGGCGCTCTCGACCGCAGCGCGCGCTTCCCGGAGAGGAACTGGAG 13246
QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
DB 13247 CACTGGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTTCTCAGCAAAAGAGCTGCC 13306
QY 221 GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
DB 13307 GCGTGTCTGGAGACAGAGGTCACTACCACTTCATCGACAGAGCTTCAGGAGTTCCTC 13366
QY 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
DB 13367 GCGGCACCTGTCTACCTGTGGAGGACGCGCGGTGCCAGGACCGCGCTGCGCGCTT 13426
QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
DB 13427 GGGACACTCTCTGGTGGGACGCCGCCACCGCACCTTGGTGTCTACCACGCGCTTC 13486
QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
DB 13487 CTCTTCGGACTGTGAGCGCGGAGCGGATGCGCGACATCGAGCGCCACTTCGCGTGCATG 13546
QY 301 ValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlnGlyCys 320
DB 13547 GTTTCAGAGCGTGTGAAGCAGGAGGCGCTGCGGTGGGTGCGAGGACAGGAGCGCTG 13606
QY 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
DB 13607 CCGGAGTGGCACAGAGGTGACCGAGGCGGCCAAAGGGCTCCGAGACACCGAGAGCCA 13666
QY 341 GluGluGluGluGluGlyGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
DB 13667 GAGGAGGAGGAGGAGGAGGAGGAGGCCCACTACCCACTGGAGTTGCTGTACTGCTGTAC 13726
QY 361 GluThrGlnGlnAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
DB 13727 GAGACGAGGAGGAGCGGTTTGTGCGCCAAAGCGCTGTGCGGTTCGCGAGCTGGCGCTG 13786
QY 381 GArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
DB 13787 CAGCGAGTGGCTTCTGCCCATGGACGTGGCTGTTCTGAGCTACTGCTGAGGTGCTGC 13846
QY 401 ProAlaGlyGlnAlaLeuArgLeuIleSerCysArgLeuValAlaAlaGlnGluLysLys 420
DB 13847 CTTGCTGACAGGACACTGGCGCTGATCAGCTGCAGATTGGTTCTGCGCAGGAGAGAGAG 13906

QY	421	LysLysSerLeuGlyLysAtsLeuGlnAlaSerLeuGlyGly	435	QY	468	-----	468
Db	13907		13966	Db	14986	GACCTAGGAAGAAGGCTGAGGGGTGACACCAGACCTGAGGCGCTGTGGGTGCTGGAGAT	15045
QY	435		435	QY	468	-----	468
Db	13967	GCAGAGATACTCTCTGTGTGTGAGGTGTGTGTGCCGGTGCAAACTGTGCACCTCAGTG	14026	Db	15046	GSCAGGCTCTGCCACACAGACAGAGGAGGAGGGGATGAGGGTCTGGGCTTCATTAC	15105
QY	435	-----	435	QY	468	-----	468
Db	14027	TCCATCTGGCATGGGGCGCACCTCCAGACACAGACCTGGGCTCCCAAGATGGGGAGGCTTT	14086	Db	15106	CAGGTTAGGTACTGTGTAGAAAAGGTTGGTCTGTGTGAGTGGAAAGCAGCTAAGCACACT	15165
QY	435	-----	435	QY	468	-----	468
Db	14087	GACCACCTCCTTGCAAGGCAACCACTGNAATCCATCTCTTCCCCAGGCTGTCCCTCACC	14146	Db	15166	GAGGAGGGATTAAGAAGACAGAGAGTGTATTTTCCCTTCTGAAGGGAGGCTGCCCTTGGA	15225
QY	435	-----	435	QY	468	-----	468
Db	14147	ATGCCACCGTCAGCCACCAACCACTATGANAACCTCCTAGTATGGCCCTACAGCTCTG	14206	Db	15226	AACTGCTGGGTTTCTCAGAAGCAGTGGAGGAAGGTGTTCAAGCGCAGGCTCCAAGGTCAGG	15285
QY	435	-----	435	QY	468	-----	468
Db	14207	GCCAGAGCCTCTTCCCATTTTACTTCTGTGAAATAAGCTCTGGAACACACAGATGCCCTC	14266	Db	15286	CTGTCTGMAATGTAGCCGGCTTCTGGTGCCTCAGTAGCAAGTCCATCTGTCCAGGCTT	15345
QY	435	-----	435	QY	468	-----	468
Db	14267	TGTCCAGAGCCTCTGCACCTTSCCCTTCCCTCCACCTAGAAATATCTCTTTGCTGCCCTGC	14326	Db	15346	CCATTTCTCCCTCTGAAATGAGCCAGTAGATCCAGAGAGTTGGGTGAGATTTAATG	15405
QY	435	-----	435	QY	468	-----	468
Db	14327	CTGCAAGTCTGCCCTTGGCTCCTCCAGCTGGCTGGGAGCCCAACTGGCTCCAGAAGTACCC	14386	Db	15406	TGGTACACATGACAGTTCTTTTTTTTTTTTTTTTTTTTGAGACGAGCCTCGCTCTGT	15465
QY	435	-----	435	QY	468	-----	468
Db	14387	TTTCGACACCTCTGTGTGAATCTGCCCCACACCTCCACAGACCTCGGCTCCTCCAG	14446	Db	15466	CGCCAGGCTGGAGTGCAGTGTGTGATCTCAGCTCAGTGCAGGCTCCGCTCCCGGGTT	15525
QY	435	-----	435	QY	468	-----	468
Db	14447	GGCAAGGCTTTTGCCAGAGTTCTTTGNAATATAGCTCCNAATTCAGCCCTGCCAGGGGC	14506	Db	15526	CACGCCATTCTCTGCCCTCAGCCTACTGAGTAGCTGGGACTACAGGTGCCCGCCACTACG	15585
QY	435	-----	435	QY	468	-----	468
Db	14507	TGGCATCTGTTGAGAGGATGAATGTACGGGTGTGSCAGGGCTGGGCTGGGGACCCA	14566	Db	15586	CCCGCTAAATGTTGTATTTTAGTAGAGACAGGGTTTTCACCGGTTAGCCAGGATAG	15645
QY	435	-----	435	QY	468	-----	468
Db	14567	AAGACCCATGTAGAGGATGTGGTGATACCGTCCCACTGACCTCACTCAGACACCCCACT	14626	Db	15646	TCTCGATCTCTGACCTCGTGATCCACCGGCTCTGCCCTCCCAAGTGTGGGATCAG	15705
QY	435	-----	435	QY	468	-----	468
Db	14627	GGGATGATGCTGGGCCACACATGATGTTGTTGCTCTGGGTTTGTTGGGGAGTGGA	14686	Db	15706	GCATGAGCCACCGCGCGGCCACATGTACAGTTCTTACATGGCAATAACTTGGCCAGTCA	15765
QY	435	-----	435	QY	468	-----	468
Db	14687	AGGGGTGGCTCAGCAAGAGGTTGAGGAGGAGGGCAGCTCTGAGACCCAGGACTACAG	14746	Db	15766	CTCTTAGCTACCAACAGTGCAGAAAGGAAGAGGAAATCTAGGAACCTAGTACTGATGG	15825
QY	435	-----	435	QY	468	-----	468
Db	14747	ATACACAGGAGTCTTTCCGCGCAGGCACTGTGAGGAGCAGGGTGGCTTCACCTTCTCT	14806	Db	15826	TGAACAGGTCCATCCTACGTTGATGGCAACTCGGCAAAATTAGCTTGATATGGGGTGGG	15885
QY	436	-----SerSerGlnGlyThrThrLysGlnLeuProAlaSerLeuLeuHisProLe	482	QY	468	-----	468
Db	14807	CTCTCCAGCAGTTCTCAAGGACACCAACAACTGCGCAGCTCCCTTCTTCATCCACT	14866	Db	15886	GGATTGGCAATCTGGAGATAAACATAAGAACCAAAAGGTGATCATTGGCCCCAGCA	15945
QY	482	upheCnAlaMetThrAspProLeuLeuHisLeuSerSerLeuThrLeu	488	QY	468	-----	468
Db	14867	CTTTCAGGCAATGACTGACCACTGTGCCATCTGAGCAGGCTCACGTG-AGTGGCCACAC	14925	Db	15946	ATTCTGCTGTTAAAGTAAAGATGTTTCAGCACACCATTTATTTAATATGAAAAAAG	16005
QY	468	-----	468	QY	468	-----	468
Db	14926	CCCGAGCTCTTCCACCGAGCGGGCTGAGACGAGGATGCCCTGGGCAGACGGAAGTAT	14985	Db	16006	GAGAAACACACAACTACCCACCACTTAGGAGATAGCGGACTAAATAAATGAGTTGGCC	16065
QY	468	-----	468	QY	468	-----	468

```

/product="Pyrafs protein"
/protein_id="AAH3139.1"
/db_xref="GI:21410402"
/db_xref="LocusID:101613"
/translation="VLKSAQTNGPRVQGEIQLRLAREGILDHKHAQFSEEDLEKLK
LRGSQVOTIFLNKKEIPGVLTKEVTVYOFIDQSFQFLAALSYLEABRTPCTPAGGVO
KLJNSDAELRGLHALITRFLGLLNTGELRDLGNHFGCVVDPHKKDITLRVQGSHP
KGPVQAKTALEIEDLEABEEEDLNFLELLYCLYETQBEDFYQALSSJFEI
VLERVLTRMDLELVNYVOCCPDGOALRLVSCGLVAAKKEKKKSLVLRKLQSST
KQPPVSLRPLCEMTTPKCHLSLILSHCLRPDAVCRDLSLEALKVAPA.RELGLLQ
SRLTNTGLSLLCEGLAWPKQVKTLMOLPDQEVNLYIVLQSPVLTLLDLSGCC
LRPVIVPEPLCAALHKPKSLKTLTSLTVELSENRLDLQAVKTSKPDLSIYSK"
BASE COUNT      884 a   909 c   902 g   759 t
CRIGIN

Alignment Scores:
Pred. No.:      2,98e-144      Length:      3454
Score:          2020.00      Matches:      4-2
Percent Similarity: 80.03%      Conservative: 65
Best Local Similarity: 69.13%      Mismatches: 201
Query Match:      62.77%      Indels:      18
DB:               10          Gaps:      8

US-10-029-347-2 (1-625) x BC03139 (1-3454)

Qy 31 AspProPheCysAlaAlaSerGlyAlaArgValLeuGlyLeuLeuSerLysAlaLeu 50
Db 31 GRCCCTTTGGAGGCTACAAAGTGGCTTTGAGAGTGGCTGAGCGGGCTGCTGAGCCAGGAGCTG 90
Qy 51 LeuProThrAlaAlaLeuLeuValThrThrArgAlaAlaAlaProGlyArgLeuGlnGly 70
Db 51 TTACCAAGAGCTCGCTTGGCTAGCTACACGCCATGCCGCTACAGGAGGGCTGCAGSGCC 150
Qy 71 ArgLeuCysSerProGlnCysAlaGluValArgGlyPheSerAspLysAspLysLys 90
Db 151 AGATTGCTCGCCACAGTGGCAGAAATACCGGGCTTCTCAGACAAAGACAAATAAAG 210
Qy 91 TyrPheTyrLysPhePheArgGluArgGluArgGluArgAlaTyrArgPheValLys 110
Db 211 TACTTCTCAAGTCTTCGCGGACGAGAGGAGCGCGCTTACCGCTTCGTGAA 270
Qy 111 GluAsnGluThrLeuPheAlaLeuCysPheValProPheValCysTrpIleValCysThr 130
Db 271 GAGAACCGAGAGCTGTTTSCACTGTGCTTTGTGCCCTTCTGCTGGATCGTGTGCACA 330
Qy 131 ValLeuArgGlnGlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThr 150
Db 331 GTGCTGCAGCAGCAGCTAGAGCTGGGCGGGATCTCTCTCGTACCTCCCAAGACCACTACT 390
Qy 151 SerValTyrLeuLeuPheIleThrSerValLeuSerSerAlaProValAlaAspGlyPro 170
Db 391 TCTGTGTACTGCTCTTCATCACCAGATGCTGTAAGTCTGCGGC---ACCAATGGACCC 447
Qy 171 ArgLeuGlnGlyAspLeuArgAsnLeuCysArgLeuAlaArgGluGlyValLeuGly--- 189
Db 448 CGGGTTCAGGGACAGCTGAGAACGCTGTGTGCGCTGCGCCAGGAGGGCATCTCGATCAT 507
Qy 190 ArgArgAlaGlnPheAlaGluLysGluLeuGluGlnLeuGluLeuArgGlySerLysVal 209
Db 508 CATAAAGACCAATTCCTCAGAAGAGGACCTGGAGAAATGAAGCTTCGGGGTTCCTCAAGTT 567
Qy 210 GlnThrLeuPheLeuSerLysLysGluLeuProGlyValLeuGluThrGluValThrTyr 229
Db 568 CAGACAAATATTCCTCAACAGAAAGAGATACCAAGGAGTGTCTAAATACTGAGGTTCACCTAC 627
Qy 230 GlnPheIleAspGlnSerPheGlnGluPheLeuAlaAlaLeuSerTyrLeuLeuGluAsp 249
Db 628 CAGTTCATTACCAAGAGCTTCCAGGAGTCTTGGCTGCATTGTCTATACCTACTAGAGCT 687
Qy 250 GlyGlyValProArgThrAlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGln 269
Db 688 GAGCGAACTCCGGGGAGCGCCCGCAGCGGGTGTGCAGAAGCTCTCTGAAGCTCTGACCGGAG 747

```

RESULT 14

```

Qy 270 ProHisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArg 289
Db 748 CTACGGTGGTCACTTTGACCTGACCAACCGATTCCTCTTTGGCTCTAAATACAGAGGGG 807
Qy 290 MetArgAspIleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAla 309
Db 808 CTTTGGTGCATTCGAAACCATTTTGGCTGTGTGCTGCCAGATCATGTGAAAGGACACC 867
Qy 310 LeuArgTrpValGlnGlnGlnGly-----GlnGlyCysProGlyValAlaProGluVal 327
Db 868 CTGCGTGGGTACAAAGACAGAGCCACCCCAAGGGGCCACCTAGTGGGGGCAAAAAGACT 927
Qy 328 ThrGluGlyAlaLysGlyLeuGluAspThrGluGluProGluGluGlnGlyGlu 347
Db 928 GCTGAG-----CTGGAGGACATCGAGAGCGAGGAGGAGGAGGAGGAAG 975
Qy 348 GluPro---AsnTyrProLeuGluLeuLeuTyrCysLeuTyrGluThrGlnGluAspAla 366
Db 976 GAGGACCTCAATTTTGGACTGGAGCTGCTGTACTGCTGTATGAGACACAGAGGAGGAT 1035
Qy 367 PheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPheCys 386
Db 1036 TTTGTTCCAGAGGCTCTCAGCAGCCTTCCAGAGATAGTACTGGAGCGAGTCAAGTTGACC 1095
Qy 387 ArgMetAspValAlaValLeuSerTyrCysValArgCysCysProAlaGlyGlnAlaLeu 406
Db 1096 CGCATGAGCCTTCAAGTTCTGAATTAATTCGCTGCAGTGTGCCCATGATGCCAGGCTCTG 1155
Qy 407 ArgLeuIleSerCysArgLeuValAlaAlaGlnGlu-----LysLysLysLysSerLeu 424
Db 1156 AGACTGGTGGCTGTGGCTGGTGGCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGCTG 1215
Qy 425 GlyLysArgLeuGlnAlaSerLeuGlyGlySerSerGlnGlyThrThrLysGlnLeu 444
Db 1216 GTGAAGCGCTGAG-----GGCTCTCAAGACCAACCAAGAACCAACCC 1257
Qy 445 ProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSer 464
Db 1258 CCAGTCTCTCTGTGCTGCCTCCTCTGTGAGACATGACTACCCCGAAATGTCATCTGAGT 1317
Qy 465 SerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerGluAla 484
Db 1318 GTTCTGATCTTGTACACTGCAGACTCCCTGTATGAGTGTTCGCGAGACCTTTCCGAGGCC 1377
Qy 485 LeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAla 504
Db 1378 CTGAAGGTAGCTCTGCGCTTAAAGGAGTTGGGCTCTCTCCAGAGCGCGGCTCACCAACACA 1437
Qy 505 GlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrValArg 524
Db 1438 GGCTGTGCTTACTGTGCGAAGGCTGCTGCTTGGCCCAAGTGCAGGTCGAGACACTCAGG 1497
Qy 525 ValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGlyMetLeuArgGlnSer 544
Db 1498 ATGCAGCTGCTGACCTCCCAAGAGGTGATCAATTACTTGGTCACTGCTGCTCCAGCAGAG 1557
Qy 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
Db 1558 CCAGTCTCTGACCACTCTGGACCTCAGTGGCTGTGAGTCTGCTGCTGCTGCTGCTGCTG 1617
Qy 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAlaSerVal 584
Db 1618 CTGTGTGCGCCCTGAGGACCCCTAATGCTCCCTAAGACCCCTCAGTCTGACCTCTGTG 1677
Qy 585 GluLeuSerGluGlnSerLeuGlnGluLeuAlaValLysArgAlaLysProAspLeu 604
Db 1678 GAGCTGAGTCAGAACTCAGTGGAGGATCTTCAAGCTGTGAAGACATCAAGGCCAGATCTG 1737
Qy 605 ValIleThrHisProAlaLeuAspGlyHisProGlnProProLysGlu 620
Db 1738 TCCATCATATAT---TCAAAATGAAGCACACCTCACCGGCTCGGAAG 1782

```

```

BC013519      2354 bp      mRNA      linear      ROD 20-SEP-2002
LOCUS      Mus musculus, clone IMAGE:4236272, mRNA, partial cds.
DEFINITION      BC013519
ACCESSION      BC013519
VERSION      BC013519.1 GI:15488764
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS      Strausberg, R.
TITLE      Submitted (04-SEP-2001) National Institutes of Health, Mammalian
JOURNAL      Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 25 Row: a Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1. .2354
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4236272"
/tissue_type="Kidney, normal, 5 month old male mouse."
/c-one_lib="NCI CGAP_Kid14"
/lab_host="DH103"
/notes="vector: pcwv-sport6"
<1..1673
/codon_start=3
/product="Unknown (protein for IMAGE:4236272)"
/protein_id="AAH13519.1"
/db_xref="GI:15488765"
/translation="LLPGARLLVTRHAATGR:QGRLCSPQCAEIRGFSKDKKKYFF
KFRPDKAERARFVKENITLALCPVPVWIVCTVQQLELGRDLSRTSKTTS
VYLLFTSMKASGTGPRVQGEIRTLRLAREGILDHKHAQFSEEDLKLKRGSOV
QTFLNKEIPGV:KTEVFYQFIDQSFQEEFLAALSYLLEAERTPGT:PAQGVQKLLNSV
AERLGHALATRE:FGLNTEGLRDIGNHFCVVPDVKKXDTLRWQGVQSHPKGPVVG
AKTALELDIEDAESEEEEDLNFGLLELYCLYETQBEDFVYQALSS:FEIVLERVR
LTMDELVINYVQCCPDGQALRVSCGLVAKEKKKKKSLVKRLKGSQSTKXQPPV
SLRPPJCTETWPKHLUSV:LSHCRLPDPAVCRDSEALKVAPALRELGLQSR:JNT
GLRLCEGLAMPK:QVKTLEMQPDQOEIVNYLVIQQSPVLTLLDLSGQCPGVIV
EPLCAALKHPKCS:KTLSLTSVLSSENSLRDLCAVKTSKPDLSIIYSK"
BASE COUNT      588 a 614 c 642 g 510 t
ORIGIN

Alignment Scores:
Pred. No.:      9.37e-140      Length:      2354
Score:          :958.00      Watches:      399
Percent Similarity: 80.24%      Conservative: 64
Best Local Similarity: 69.15%      Mismatches: 96
Query Match:      60.85%      Indels:      18
DB:              10      Gaps:      8

```

```

US-10-029-347-2 (1-625) x BC013519 (1-2354)
QY      50      LeuLeuProThrAlaLeuLeuLeuValThrThrArgAlaAlaProGlyArgLeuGln 69
DB      3      CTGTTACCAGGAGCTCGCTTGTAGTACTACAGCCATCGCGCTACAGGAGGCTGCAG 62
QY      70      GlyArgLeuCysSerProGlnCysAlaGluValArgGlyPheSerAspLysAspLysLys 89
DB      63      GGCAGATGTGCTGCCACAGTCCGAGAAATACGCGGCTTCTCAGACAAACAAAAA 122
QY      90      LysTyrPheTyrLysPhePheArgAspGluValArgAlaGluArgAlaTyrArgPheVal 109
DB      123      AAGTATTTCTTCAAGTTCTTCGGCAGCAGAGAGGAGGAGCGCTTACCGCTTCGTG 182
QY      110      LysGluAsnGluThrLeuPheAlaLeuCysPheValProPheValCysTyrPheValCys 129
DB      183      AAAGAAACGACAGACCTGTTTGCATCTTGTGCTTGTGCTTGTGCTGCTGCTGCTGCTG 242
QY      130      ThrValLeuArgGlnGlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThr 149
DB      243      ACAGTGCTGCAGCAGCAGCTAGAGCTGGGCCGGATCTCTCTGTTACCTCCAGACCACT 302
QY      150      ThrSerValTyrLeuLeuPheThrSerValLeuSerSerAlaProValAlaAspGly 169
DB      303      ACTTCTGTGTACTGCTCTTCTATCACCAGCATGCTGAAGTCTGCAGGC---ACCAATGGA 359
QY      170      ProArgLeuGlnGlyAspLeuArgAsnLeuCysArgLeuAlaArgGluGlyValLeuGly 189
DB      360      CCCCCTGTTTACGGGAGAGCTGAGAACGCTGTGTCTGCTGGCCAGGAGGAGGAGGAGGAT 419
QY      190      ---ArgArgAlaGlnPheAlaGluLysGluLeuGlnLeuGluLeuArgGlySerLys 208
DB      420      CATCATAAAGCACAATCTCAGAGAGGAGCTGGAGAAATGGAAGCTTCGGGGTTCCCAA 479
QY      209      ValGlnThrLeuPheLeuSerLysGluLeuProGlyValLeuGluThrGluValThr 228
DB      480      GTTCAGACAATATTTCTCAACAAGAAAGAGATACCAGGAGTCTTAAAACTCAGGTCAACC 539
QY      229      TyrGlnPheLeuAspGlnSerPheGlnGluPheLeuAlaLeuSerTyrLeuLeuGlu 248
DB      540      TACCAGTTCATTGACCAGAGCTTCCAGGAGTTCCTTGGCTGCAATTCATACCTACTAGAA 599
QY      249      AspGlyGlyValProArgThrAlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAla 268
DB      600      GCTGAGCGNACTCCGGGAGCGCCGAGGGGTGTGCAGAACCTCTGNACTCTGACGCG 659
QY      269      GlnProHisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGlu 288
DB      660      GAGTACGTGGTCACTTGCACCTGACCCCGCATTCCTTTGGCTGTGCTGCTGCTGCTGCTG 719
QY      289      ArgMetArgAspLeuGluArgHisPheGlyCysMetValSerGluArgValLysGlnGlu 308
DB      720      GGGCTTCGTGACATGGAAACCATTTTGGCTGTGGTGGCCAGATCATGTGAAAGAGGAC 779
QY      309      AlaLeuArgTyrValGlyGlyGlnGly-----GlnGlyCysProGlyValAlaProGlu 326
DB      780      ACCCTCGGTGGGTACAGGACAGACCCACCAAGGGGCGCCACCATAGTGGGCGCAAAAAAG 839
QY      327      ValThrGluGlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGluGluGly 346
DB      840      ACTGCTGAG-----CTGAGGACATCGAGGACGAGGAGGAGGAGGAGGAGGAGGAA 887
QY      347      GluGluPro---AsnTyrProLeuGluLeuLeuTyrCysLeuTyrGluThrGlnGluAsp 365
DB      888      GAGGAGGACCTCAATTTTGGAGCTGGAGCTGTGCTACTGCTGTATGAGACACAGAGGAG 947
QY      366      AlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPhe 385
DB      948      GATTTTGTTCAGAGGCTCTCAGAGCCCTTCAGAGATAGTACTGGAGCGAGTCAAGTTG 1007
QY      386      CysArgMetAspValAlaValLeuSerTyrCysVala:gcysCysProAlaGlyGlnAla 405
DB      1008      ACCCGCATGGACCTTGAAGTTCTGAATTTACTGCGTGCAGTGTGCCAGATGCCCGAGCT 1067

```


Consensus quality: 225463 bases at least Q30
 Consensus quality: 226655 bases at least Q20
 Estimated insert size: 230654; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length.
 (see http://www.hgsc.bcm.tmc.edu/docs/GenBank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 4 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

* 1 222461: contig of 222461 bp in length
 * 222462 222561: gap of unknown length
 * 222562 224949: contig of 2388 bp in length
 * 224950 225049: gap of unknown length
 * 225050 228242: contig of 3193 bp in length
 * 228243 228342: gap of unknown length
 * 228343 239533: contig of 10191 bp in length.

FEATURES
 source
 1. 239533
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-172C3"

misc_feature
 1. 2349
 /note="wgs contig"

misc_feature
 124155..239379
 /note="clone boundary
 clone_end:T7
 site:
 end_sequence:3H290527"

BASE COUNT 61451 a 53521 c 52385 g 60005 t 1147: others
 ORIGIN

Alignment Scores:
 Pred. No.: 4,63e-121 Length: 238533
 Score: 1745.50 Matches: 452
 Percent Similarity: 36.80% Conservative: 57
 Best Local Similarity: 32.68% Mismatches: 113
 Query Match: 54.27% Indels: 763
 DB: 2 Gaps: 11

US-10-029-347-2 (1-625) x AC132968 (1-238533)

Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
 Db 170887 ATACTAGCGCAACCGCATCGCTACTCTGACGAGCGGTGCAGACGAATTCGCTACT 170946

Qy 21 LeuGlyGlyProGluAlaAlaProCysThrAspPheGluAlaAlaSerGlyAlaArg 40
 Db 170947 TTACAGCCCCAGAAAGCCATCCCTCGCAGAGACCTTCGAGGCTACGAGTGGCTTGAGA 171306

Qy 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 60
 Db 171007 GTGCTGACGGGCTGCTGAGCAGAGAGCTGTACCATCGGCTCGCTTGGTGGCTTCA 171656

Qy 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
 Db 171067 CGCAACGCCACCCCTCGGAGGCTCGAGGGCAGACTTTGCTCACCACAGTGCAGGAGAAGTG 171126

Qy 81 ArgGlyPheSerAspLysLysLysLysLysLysLysLysLysLysLysLysLysLys 100
 Db 171127 CGCGCTTCTCAGACAAAGACAGAAAGATTTCTTCAAGTTTTCCTCGGAGCAGAGG 171186

Qy 101 ArgAlaGluArgAlaTyArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
 Db 171187 AAGCAGAGCGCGCTTACCGCTTGTGMAAGAGATGAGACGCTGTACGCACTGCTGCTTT 171246

Qy 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140

Db 171247 GTGCGGTTTGTGCTGCTGATCGTGTGCACCGTGTGTGCAGCAATGAGCTGGCGCGG 171306

Qy 141 AspLeuSerArgThrSerLysThrThrSerValThrLeuLeuPheIleThrSerVal 160
 Db 171307 GATCTGTCTCGTACCTCAAGACCACTACATCTGTACCTGCTCTTATCACCAGCATG 171366

Qy 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
 Db 171367 CTGAAGTCTGCAGGT--ACCAATGGACCCCGGTTACAGGAGAGCTGCGAATGCTGTGC 171423

Qy 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGlnLysGluLeuGlu 200
 Db 171424 CGCCTGGCGCGGAGGCGATCCTGAAGCATCAAGCACAGTTCTCAGAAAGACCTGGAG 171483

Qy 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
 Db 171484 AGATTGAAGCTTTCAGGGTTCCCAAGTTCAGACAAATGTTTCTCAGCAAGAGGCTGCCA 171543

Qy 221 GlyValLeuGluThrGluValThrTyArgPheIleAspGlnSerPheGlnGluPheLeu 240
 Db 171544 GGAGTCTAGAAACTGTGTGTCACCTACCAGTTCAATTGACAGAGCTTCCAGAGTCTTG 171603

Qy 241 AlaAlaLeuSerTyLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
 Db 171604 GGTGCTATGTCTACCTACTAGACGCTGAGGAGGCCCCAGGGAACCTCCGACAGAAATGTG 171663

Qy 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
 Db 171664 CAGATGCTCTGAATCTCTGCGGGGCTGCTGTGTCATCTGCGCACTCACCCTAGATTTC 171723

Qy 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspLeuGluArgHisPheGlyCysMet 300
 Db 171724 CTCCTTTGACTCTAAGTACAGAGAGGATTCGTGACATTTGGAACCATTTTGGCTGTGTG 171783

Qy 301 ValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlnGlyGlnGlyCys 320
 Db 171784 GTGCCAGGCGCTGTGAACAGGACACCTTGGCGTGGGTACAAGGCAAAAGCCAA----- 171837

Qy 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAsp----- 336
 Db 171838 CCCAAGGTGGCGACAGTAGGACGAGAAAGAGGATGAGCTGAGGACGAGGAGGAGAG 171897

Qy 337 ThrGluGluProGluGluGluGluGluGluGluProAsnTyProLeuGluLeuLeu 356
 Db 171898 GAGG 171957

Qy 357 TyrCysLeuTyThrGluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPhePro 376
 Db 171958 TACTGCTCTATGAGACACAGAGGATGATTTGTTCGCGAGGCTCTCAGCAGCCTTCCA 172017

Qy 377 GluLeuAlaLeuGlnArgValArgPheCysArgMetAspValAlaValLeuSerTyCys 396
 Db 172018 GAGATGGTACTGGAGCGAGTTAGGCTGACCCGATGGACCTTGAGGTTCTGAGCTACTGC 172077

Qy 397 ValArgCysCysProAlaGlyGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAla 416
 Db 172078 GTGCAGTGTGCGCGAGCGGCGAGGCTCTGAGACTGGTGAGCTGTGAGTGTGGTGGCGGCA 172137

Qy 417 GlnGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 422
 Db 172138 AAGGAGAAAGAAAGAAAGAAAGAGAGCTTCATGAACCGGCTGAAGGGGTGAGTCTAGAG 172197

Qy 422 ----- 422
 Db 172198 GTGAGGTCTCTCGCCGATGCTACTGTGCCACACAGCCCATGCCACTCGGTTGAGAA 172257

Qy 422 ----- 422
 Db 172258 GCATTCCACATCTGGTAAACCAGATGGCAGAGCGGACCTCAGAGCTGCTTACCCATCCC 172317

Qy 422 ----- 422

Db 172318 TTGCAAGGAGCGCTCTCTATTCTCTTCAAAATGACCTCACCACGAGTACTCTCAGTCAAT 172377
Qy 423 -----SerLeuGlyLysArgLeuGlnAlaSer--- 431
Db 172378 GCGATCAAGGCNACCTCCNAGTTCAGCCAGGTAGGTCCAGGAG-CNACCGAGCCCC 172436
Qy 431 ----- 431
Db 172437 CTCTCTGATGCACTTCTCATACATTTCTGGAATATGCCATGTCTCTGTGTCTGAGCCT 172496
Qy 431 ----- 431
Db 172497 TGTACCTGGGGCTCCCTCCATCCAAACACTATTTCGTCTCTCTTCATATGCTACTCTG 172556
Qy 431 ----- 431
Db 172557 COTCCACCACCTTGGGTCCAACTACTGCCCCCAAGGCACCCCAATTGACCACATTATTCTC 172616
Qy 431 ----- 431
Db 172617 ATATCTTCTTCTTGGCCCACTCGTCCACCTGATCCTAGGTTTCCACGACATGGGCAATTTGTT 172676
Qy 431 ----- 431
Db 172677 TTTTATGAGCTCAGCTCTACACATGCTTGTCCCAAGGCTCGCCCACTGTTGAGAGTATA 172736
Qy 431 ----- 431
Db 172737 CACCAGGATAATGGCAGGACCTGCTCTTGGGATCTGCAGCCCCACACTTGGGTTTGTCTC 172796
Qy 432 -----LeuGlyGly--- 434
Db 172797 AAATTGGTTGGGGAGAGGTGCAAGGATGTGAGATAAATGATGTTCTTGGAACTCAGG 172856
Qy 434 ----- 434
Db 172857 AAGGATTGAGGCTTTCCAGAAAGGCTGGNAGGGGTGCTACAGAGCTCAACCTTGTCTC 172916
Qy 435 -----GlySerSerGlnGlyThrLysGlnLeuProAlaSerLeuLeuHisPro 451
Db 172917 TTCTCTCCCTCCAGTTCTCAAGACCGGGGAACAACCCCGACCTCCCTTGTGTGGTCCA 172976
Qy 452 LeuPheGlnAlaMet----- 456
Db 172977 CTCTGTGAGGCATGATTACCCAGCAATGTGGTCTGAGTATCTGACGTGAGTCACCACTG 173036
Qy 456 ----- 456
Db 173037 CTCTGGAGGGGATTTACACTGGGTGGTCTGACAGAGGACAGAAATTTGCTTCTAAGGAA 173096
Qy 456 ----- 456
Db 173097 GGGTAAGGACATAGGCGCTACTGGCTTTTCAACCACTATTTGGGGCTGTGACGGTGGGGCTG 173156
Qy 456 ----- 456
Db 173157 GTGCNAGACCTTGTGGGAATCAGAAAGCTGCACTGCTTTGGAGACGGCGCAGAGAAGCA 173216
Qy 456 ----- 456
Db 173217 ACAAGTGTTCCTTTGAAGACAGCTCTTAAGTGTGTAAGGACACTGCCCCGCCACCCCT 173276
Qy 457 -----ThrAspProLeuCysHisLeu--- 463
Db 173277 ATGCAGAAAGTGTGAGGGAGCTCAGAGCTGGTGCACCCCTGCTCTGTGTGCTTGA 173336
Qy 463 ----- 463
Db 173337 GCAAGTCAGTCTGCCCATGCACTGGTGTCTCTCCCTGTAAACATACCCAGCTGATCCAGCT 173396
Qy 463 ----- 463
Db 173397 CACAGGATTAGTGAAGATTATCTGAGGCTCCACGGAGAAGTCTTCTTAAGATATGCTAGG 173456

Qy 463 ----- 463
Db 173457 TTTTCTTATTACTGTGAGCTATCTCTCAGTGAGCGAAGAGTAAAGACAAATTTACAGGAACCAA 173516
Qy 463 ----- 463
Db 173517 TAGAGCAGGCCATGTTCGAAGCAGACAGTCAAGGATGATGCAATTGATAGAGGATGAGGAG 173576
Qy 463 ----- 463
Db 173577 GAAAACTGGCTCTGTGGGGGTGAGGCCCCAGCAACGTGAAAAATAAAACAACGGAGCCGA 173636
Qy 463 ----- 463
Db 173637 AAAGGCAGTCGATTTGGGCTCAGGAATTTCTACGAAAAACCAAGTATACAGAGGTGCTCA 173696
Qy 463 ----- 463
Db 173697 CTTTCATAATTTGTAATGGGAAAAGATAAATAATCAGAAATAGCAGAGCTGTTTCAGCCAC 173756
Qy 463 ----- 463
Db 173757 TGAGGATGACTGCTCAGCCACTGAGGATGACTGCTCAGCCACTGGGTATGACTGCTCAGC 173816
Qy 463 ----- 463
Db 173817 CACTGGGGATGACTGCTCAGCCACTAGGGGATGACTGCTCAGCCACAGGGATGACTGCTC 173876
Qy 463 ----- 463
Db 173877 AGCCACTGGGGATGACTGCTCAGCCACTTGGGGATGACTGACAGCCCATCCGAACCAAAAT 173936
Qy 463 ----- 463
Db 173937 ATGAAGAACTGTAGATTTCTTCACTGAACATCAGTCTGCCAGTGAATAGTATGTTTGTTC 173996
Qy 464 -----SerSerLeuThr-----LeuSe 469
Db 173997 AGCTGAGGTAGAGAGTGGCCAGTCTTACAGTCTCCATGCTCTTTCAGCCAGCTGTCTC 174056
Qy 469 rHisCysLysLeuProAspAlaValCysArgAspLeuSerGlnAlaLeuArgAlaAlaPr 489
Db 174057 ACACGTCAAACTCCCTGATGCACTTTGTGAGACCTTTCTGAGGCTCTCAAGGTAGCTCC 174116
Qy 489 oAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLe 509
Db 174117 TCCCTAAGGGAGCTGGGCTCTCCAGAAACCGACTCACTGAGGGCAGGCTGCGTTTACT 174176
Qy 509 uSerGluGlyLeuAlaLysProGlnCysArgValGlnThrVal----- 523
Db 174177 AAGCCAAAGGCTGGCTTGGCCCCAAATGCAAGGTGCAACACTCAGGTAAAGGCTGGTTT 174236
Qy 524 -----Ar 524
Db 174237 GGAAGGTGGGGGTAGGGAACCATCCCAACAGCAACCCCTGAGTTTGGCTTTCCCCACAG 174296
Qy 524 gValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGlyMetLeuArgGlnSe 544
Db 174297 GATACAGATGCTGGGCTCCAGAGGTGATCCACTACCTGGTCTGTCATTGTGCTCAGCAGAG 174356
Qy 544 rProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTy 564
Db 174357 CCCAGCTTAACCACTCTGACCTCAGTGGCTGTCTCAGTCTGCTGGCCTGCTGGTGGNAOC 174416
Qy 564 rLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeu----- 579
Db 174417 TCTGTGTTTCAAGCCCTGAAGCACCCCTAAATGTGGCCTTAAAGACCCCT-CAGGTAAAGCAGG 174475
Qy 579 ----- 579
Db 174476 GGGACGCAATGGTCTCTGAGGACAGAGCCAGTGAACCCCTGATGGGAGAGGTGCTTTGGCG 174535


```
QY 579 ----- 579
Db 174536 GCGTCTCTGTGAGACCTCACACAGATCACACAGATCATGATTGAGCTCACAGCTCTG 174595
QY 579 ----- 579
Db 174596 GGTGACGAGCCCATTAAGACATACCTTCTGAGGCCCTATCCAAAGCTCCTGTCTCAAA 174655
QY 579 ----- 579
Db 174656 CAAAGGGCCCTAAGCTGTGGGCACAGTCTGAGCATAGCTCTCTCTGTATCCCTCAGAG 174715
QY 579 ----- 579
Db 174716 GGCCTCAGCTTTTAATAGACCTCAACCGACCTGCCCTCTGTACTTGTATATCTAAGCA 174775
QY 579 ----- 579
Db 174776 AGCAGTGCCAAACACCCCGCTCTCACCTCAAGCCCTGGGTTCTCCCCACAGGCAGTCGGC 174835
QY 580 -----Ser-LeuAlaSe 583
Db 174836 CTTTGGCCCTGACTGATCCTTGTTCGCCCTGTCTTACACATGTGGCTGCAGTCTGACTTC 174895
QY 583 rVa:GluLeuSerGluGlnSerLeuGlnGluLeuGlnAlaValIysArgAlaIysProAs 603
Db 174896 TGTGAGCTGACTGAGAATCTCACTGAGAGAGCTTCAAGCTGTGAAGACATTAAAGCCAGA 174955
QY 603 pLeuValIleThrHisPrcaIalaLeuAspGlyHisProGlnProProIysGluLeuIleSe 623
Db 174956 TCTGGCCATCATACATTCAAAATTGGGCACACATCTCTCAGCCCTCTGAAGGGATGAAGCAG 175015
QY 623 rThrPhe 625
Db 175016 TGTCTTC 175022
```

Search completed: October 3, 2003, 15:49:12
Job time : 6649 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 3, 2003, 09:15:40 : Search time 442 Seconds
(without alignments)
3877.079 Million cell updates/sec

Title: US-10-029-347-2
Perfect score: 3218
Sequence: 1 MAAQQRLLFLDCAADLPA.....ITHPALDGHCPKPELISTF 625

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n_model_DEV=xih
-Q=/Cgn2_1/USPFO/spool/US10029347/runat_02102003_173436_17066/app_query.fasta_1.775
-DB=N Geneseq_195Jun03 -QEM=fastcap -SUFFIX=trg -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029347 @Cgn_1_1_312 @runat_02102003_173436_17066 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_195Jun03:
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3218	100.0	2689	24	ABS63485	DNA encoding human
2	3218	100.0	3365	22	AAD08836	Human G-protein co
3	3170	98.5	2763	24	ABS63486	Human leucine-rich
4	3144	97.7	2109	24	ABX97042	Human NOV18a cDNA.
5	2882	89.6	5556	24	AAL47132	Pyrin domain conta
6	2882	89.6	6461	22	AAS00033	Human ATAS-2-enco
7	2239	69.6	1795	22	ABA08979	Human vasopressin
8	2209	68.6	2054	24	ABS63487	Human leucine-rich
9	1737	54.0	2251	14	AAQ38723	ATP/ADP2 receptor
10	1243	38.6	933	22	AAK83313	Human immune/haema
11	1167.5	36.3	989	22	AAK62741	Human immune/haema
12	1147.5	35.7	990	24	ABL90582	Human polynucleoti
13	826.5	25.7	3300	24	AAL47129	Pyrin domain conta
14	825.5	25.7	3186	24	AAL44363	Human PYRIN-8 cDNA
15	800.5	24.9	3857	22	AAD14323	Human PYRIN 1 (PYR
16	800.5	24.9	3857	25	ABX93556	Human cDNA encoding
17	791	24.6	704	24	ABQ29972	Oligonucleotide fo
18	791	24.6	704	24	ABQ29972	Oligonucleotide fo
19	789	24.5	704	24	ABQ29970	Oligonucleotide fo
20	789	24.5	704	24	ABQ29971	Oligonucleotide fo
21	762	23.7	4422	22	AAD02760	Human NB-ARC and C
22	762	23.7	4556	22	AAD02764	Human NAC beta iso
23	762	23.7	5122	22	AAD12951	Human G-protein co
24	762	23.7	5523	22	AAF83651	Human CARD-7 poly
25	762	23.7	5523	25	ABS55497	cDNA encoding huma
26	762	23.7	5523	22	ABS56030	cDNA encoding huma
27	760.5	23.6	2847	22	ABS01487	Human secreted pro
28	760.5	23.6	2847	25	ABZ73494	Secreted protein-e
29	758	23.6	5100	24	AAL47127	Pyrin domain conta
30	716	22.2	4466	22	AAD02765	Human NAC gamma or
31	712.5	22.1	4194	22	AAD02761	Human NB-ARC and C
32	712.5	22.1	4329	22	AAD02762	Human NB-ARC and C
33	694	21.6	4931	24	ABL59333	Nucleotide sequenc
34	691	21.5	2763	24	ABQ75801	Human MDDT-13 enco
35	680.5	21.1	678	25	ACA56413	Norway rat signal
36	660	20.5	3085	23	AAS68111	DNA encoding rove
37	650	20.2	3263	24	AAI70684	Human nucleotide b
38	650	20.2	3459	24	ABQ78049	Human CGPD encodi
39	635	19.7	3069	24	AAL47134	Pyrin domain conta
40	634	19.7	3150	22	AAI67184	Nucleotide sequenc
41	619.5	18.3	3431	22	AAD14322	Human nucleotide b
42	619.5	18.3	3431	25	ABX93555	Human cDNA encoding
43	616.5	18.2	3189	22	AAI67185	Nucleotide sequenc
44	616.5	18.2	3531	24	AAL47128	Pyrin domain conta
45	603.5	18.8	2218	20	AAZ20641	CBDAK001 coding se

ALIGNMENTS

RESULT 1:
ABS63485
ID ABS63485 standard, cDNA, 2689 BP.
XX
AC ABS63485;
XX
DT 15-NOV-2002 (first entry)
XX
DE DNA encoding human leucine-rich repeat small intestine I (HLRRS11).
XX
KW Human; human leucine-rich repeat small intestine I; HLRRS11; asthma;
KW proliferative disorder; gastrointestinal disorder; renal disorder;
KW neural disorder; reproductive disorder; calcium regulation; apoptosis;
KW immune system; anaemia; human immune deficiency virus; HIV; cancer;
KW blood coagulation disorder; autoimmune disorder; allergic reaction;
KW inflammatory condition; cardiovascular disorder; ischaemia;
KW neurological disorder; infectious disease; cytokine production;

expressed sequence tag; EST; gene; ss.

KW expressed sequence tag; EST; gene; ss.
 XX Homo sapiens.
 OS WO200261086-A2.
 PN 08-AUG-2002.
 PD 20-DEC-2001; 2001-WO-US49739.
 PF 22-DEC-2000; 2000US-257774P.
 PR (BRIM) BR-STOL-NYERS SQUIBB CO.
 PA
 XX Feder J, Ramanathan C, Mintier G;
 XX WPI; 2002-619252/66.
 DR P-PSDB; ABG78454.
 XX New isolated nucleic acid molecules encoding HLRRS11 polypeptides, or
 PT their fragments and homologues, useful for preventing, treating and
 PT ameliorating medical conditions, e.g. proliferative, gastrointestinal,
 PT or renal disorders.
 XX
 PS Claim 1; Figure 1; 336pp; English.
 XX The invention relates to isolated nucleic acid molecules (1) encoding
 CC human leucine-rich repeat small intestine I (HLRRS11) polypeptides.
 CC The nucleic acid molecules and polypeptides are useful for preventing,
 CC treating and ameliorating medical conditions, such as proliferative,
 CC gastrointestinal, renal, neural, or reproductive disorders; or disorders
 CC related to aberrant calcium regulation or apoptosis modulation; either
 CC directly or indirectly. They are also useful for treating, preventing
 CC and/or diagnosing diseases, disorders and/or conditions of: immune system
 CC mobilization or inhibiting the proliferation, differentiation, or
 CC maturation of immune cells; haematopoietic cells e.g. thrombocytopenia,
 CC anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
 CC virus (HIV) infection, HTLV-ELV infection; blood coagulation disorders,
 CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
 CC myasthenia gravis; asthma or allergic reactions; inflammatory
 CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
 CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
 CC ischaemias, aneurysms; neurological disorders, e.g. Alzheimer's disease,
 CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
 CC pneumonia, or viral, bacterial, and fungal infections. The HLRRS11
 CC polypeptides are useful for modulating cytokine production, antigen
 CC presentation, or other processes such as boosting immune responses.
 CC ABS63485-ABS63504 represent HLRRS11 coding sequences and PCR primers of
 CC the invention.
 XX
 SQ Sequence 2689 B?; 534 A; 961 C; 844 G; 450 T; 0 other:

Alignment Scores:

Pred. No.: 1,1e-269 Length: 2689
 Score: 3218.00 Matches: 525
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-10-029-347-2 (1-625) x ABS63485 (1-2689)

QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuLeuAspGlyValAaAspGluLeuProAla 20
 DB 75 ATCTGGCCCGCCAGCCGGCTGCTTCATCTCTGGAGCGCGCGGACGACTGCGCGCG 134
 QY 21 LeuGlyGlyProGluAlaAaProCysThrAspProPheGluAlaAaSerGlyAlaArg 40
 DB 135 CTGGGGGGCCCGAGCCCGCCCTGCACAGACCCCTTCGAGCGCGCGAGCGCGCGG 194
 QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
 DB 195 GTGTAGGCGGGCTGCTGAGTAAAGCGGCTGCTGCCACCGGCTCTCTCTGTCACACG 254

QY 61 ArgAaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
 DB 255 CGCGCGCGCGCCCGGAGGCTGCAGGCGCGCTGTTCCCGCAGTGCAGCGAGGTG 314
 QY 81 ArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLysLys 100
 DB 315 CGCGGCTTCTCCGACAAAGGACAAGAAAGATATTCTTCAAAGTCTTCCGGAAGAGAGG 374
 QY 101 ArgAlaGluArgAlaTyrArgPheValLysGluAenGluThrLeuPheAlaLeuCysPhe 120
 DB 375 AGGCGCGAGCGGCTACCGCTTCTGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 434
 QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
 DB 435 GTGCCCTTCTGTTGCTGATGCTGTGCACCGTGTGCAGCGAGAGAGAGAGAGAGAGAG 494
 QY 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
 DB 495 GACCTGTCCGCGACGCTCCAGACACACAGCTGAGTGTACCTGTCTTTCATCACCAGCGT 554
 QY 161 LeuSerSerAlaProValAlaAaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
 DB 555 CTGAGCTCGGCTCCGCTAGCGAGCGGCGCGGTTGCAGGGCGAGCTCGCGCAATCTGTGC 614
 QY 181 ArgLeuAlaArgGluGlyValLeuGlyArgAlaGlnPheAlaGluLysGluLeuGlu 200
 DB 615 CGCCTGGCCCGCGAGGGGCTCTCGAGCGAGCGGCGAGTTCGCGAGAGAGAACTGGAG 674
 QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
 DB 675 CAACTGGAGCTTCGTGGCTCCAAAGTGCAGAGCGCTGTTCTCAGCAAAAGAGAGCTGCGG 734
 QY 221 GlyValLeuGluThrGluValThrTyrGlnPheLeuAspGlnSerPheGlnGluPheLeu 240
 DB 735 GSGCGTCTGGAGACAGAGTCACTTACCAGTTCATCGACAGAGAGTTCAGAGAGTTCCTC 794
 QY 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
 DB 795 GCGGCACTGCTCACTCTCTGTGAGGAGCGGGGGTGCAGGAGCGCGGCTTGGCGGCGCT 854
 QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
 DB 855 GSGACACTCTCGTGGGAGCGCCCGAGCGGACACCCACTTTGGTGCTCACCACGCGCTTC 914
 QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
 DB 915 CTCTTCGGACTGCTGAGCGCGGAGCGGATGCGGACATCGAGCGGCACTTCGGCTGCAATG 974
 QY 301 ValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlyGlyGlyCys 320
 DB 975 GTTTCAGAGCGTGTGAGAGCAGGAGGCGCTTGGTGGTGCAGGAGAGAGAGAGAGAGAG 1034
 QY 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
 DB 1035 CGCGAGTGGCACAGAGGTGACCGAGGGGGGCAAGAGGCTCGAGAGACACCGAGAGAGCA 1094
 QY 341 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 360
 DB 1095 GAG 1154
 QY 361 GluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
 DB 1155 GAGAGCGAGAGAGAGCGCTTGTGCGCCAAAGCGCTTGTCCCGTTCCTCCGAGCTGGCGCTG 1214
 QY 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
 DB 1215 CAGCGAGTGGCTTCTGCGGCAATGGAGCTGGCTGTCTTCTGAGCTACTGCTGAGGTGCTGC 1274
 QY 401 ProAlaGlyGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluLysLys 420
 DB 1275 CTTGCTGGACAGGCACTGCGGCTGTATCAGCTGCAGATGGTGTGCTGCGCAGAGAGAGAG 1334

	QY	421	LysGlySerLeuGlyLysArgCLeuGlnAlaSerLeuGlyGlyGlySerGlnGlyThr	440
	DB	1335	NAGAAGAGCGCTGGGGGAAGCGCCCTCCAGGCAGCGCTGGGTGGCGGCAgTCTCAAGGCACC	1394
	QY	441	ThrLysGlnLeuProAlaSerLeuHisProLeuPheGlnAlaMetThrAspProLeu	460
	DB	1395	ACAAACAACCTGCCAGCGCTCCCTTCATCCACTCTTTTCAGGCAATGACTGACCCCACTG	1454
	QY	461	CysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAsp	480
	DB	1455	TGGCACTGAGCAGCGCTCAAGCTGTCCCATCGCAACTCCCTCAGCGGGTCTGCCGAGAC	1514
	QY	481	LeuSerGluAlaLeuArgAalaProAlaLeuThrGlnLeuGlyLeuLeuHisAsnArg	500
	DB	1515	CTTCTTGAGGCCCTCAGGGGAGCCCCCGCACTGACGGAGCTGGCGCTCTCTCCACAACAGS	1574
	QY	501	LeuSerGluAlaGlyIleuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgVal	520
	DB	1575	CTCAGTGAGCGGGAGCTGGGTATGCTGAGTGAGGGGCTTACGCTTGCGCGCAGTGCAGGGTG	1634
	QY	521	GlnThrValArgValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGlyMet	540
	DB	1635	CAGACCGTTCAGGTACAGCTCCCTGACCCCAGGAGGGGCTCCAGTACCTGTGTGGGTATG	1694
	QY	541	LeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaPro	560
	DB	1695	CTTGGGCAGAGCCCCGCCCTGACCACTCGAATTCTCAGCGGCTGCCAACCTGCCCGCCCC	1754
	QY	561	MetValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSer	580
	DB	1755	ATGGTGACCTACCTGTGTGGCAGTCTCTGCACACCCAGGGATGGCGCTCTGCAGACCCCTCAGT	1814
	QY	581	LeuAlaSerValGluLeuSerGluGlnSerLeuGlnGluLeuGlnAlaValLysArgAla	600
	DB	1815	CTGGCCCTCTGTGAGCTGACGGAGCAGTCACTACAGGAGCTTCAGGCTGTGAGAGACGA	1874
	QY	601	LysProAspLeuValIleThrHisProAlaLeuAspGlyHisProGlnProproLysGlu	620
	DB	1875	AAGCGGATCTGTCTATCACACACCCAGCGCTGGAGCGCCACCCACAACCTCCCAAGGAA	1934
	QY	621	LeuIleSerThrPhe	625
	DB	1935	CTCATCTCGACCTTC	1949
	RESULT 2			
	AAD08836			
	ID	AAD08836 standard; cDNA; 3365 BP.		
	XX	AAD08836;		
	AC			
	XX			
	DT			
	XX			
	DE	04-SEP-2001 (first entry)		
	XX	Human G-protein coupled receptor-2 (GREC-2) cDNA.		
	DE			
	XX	Human; G-protein coupled receptor-2; GREC-2; gene therapy; cirrhosis;		
	KW	transgenic animal; proliferative disorder; actinic keratosis; hepatitis		
	KW	nephrotropic; cancer; breast; bladder; bone marrow; brain; uterus;		
	KW	leukemia; adenocarcinoma; lymphoma; melanoma; myeloma; epilepsy; stroke;		
	KW	neurological disorder; Alzheimer's disease; Parkinson's disease; nausea;		
	KW	Huntington's disease; multiple sclerosis; dementia; angina pectoris;		
	KW	central nervous system disorder; cardiovascular disorder; hypertension;		
	KW	atherosclerosis; congestive heart failure; gastrointestinal disorder;		
	KW	dysphagia; peptic esophagitis; spasm; gastritis; anorexia; pyrosis;		
	KW	pancreatitis; Crohn's disease; diarrhoea; autoimmune disorder; anaemia;		
	KW	inflammatory disorder; Acquired Immune Deficiency Syndrome; AIDS;		
	KW	Addison's diseases; allergy; asthma; diabetes mellitus; antithyroid;		
	KW	atopic dermatitis; glomerulonephritis; Grave's disease; psoriasis;		
	KW	rheumatoid arthritis; ulcerative colitis; osteoporosis; antifungal;		
	KW	metabolic disorder; obesity; nontropic; protozoacide; virucide; ss.		
	OS	Homo sapiens.		
	XX			
	FF	Key	Location/Qualifiers	
	FX			

FT	CDS	1..2673
FT	/FT	/*tag= a
FT	/product=	"Human GCRC-2 protein"
XX		
PN	WO200142288-A2.	
XX		
PD	14-JUN-2001.	
XX		
PP	07-DEC-2003; 2000WO-US33382.	
XX		
PR	10-DEC-1999; 98US-0172852.	/
PR	22-DEC-1999; 99US-0171732.	
PR	14-JAN-2000; 2000US-0176148.	
PR	21-JAN-2000; 2000US-0177331.	
XX		
PA	(INCY)- INCYTE GENOXICS INC.	
XX		
Pf	Burford N, Baughn MR, Au-Young C, Yang J, Lu DAM, Reddy R;	
XX		
DR	WPI; 2001-381635/40.	
XX	P-PSDB; AAE04546.	
DR		
XX	New human G-protein coupled receptor polypeptides for diagnosing,	
PT	preventing, and treating cell proliferative, neurological,	
PT	cardiovascular, gastrointestinal, autoimmune and metabolic disorders -	
XX		
PS	Claim 5; Page 159-160; 175pp; English.	
XX		
CC	The present sequence is human G-protein coupled receptor-2 (GCRC-2)	
CC	cDNA. GCRC is useful in somatic or germline gene therapy to correct a	
CC	genetic deficiency, to express a conditionally lethal gene product and	
CC	to express a protein which affords protection against intracellular	
CC	parasites and also for diagnosis of disorders associated with expression	
CC	of GCRC. GCRC is also useful for generating hybridisation probes useful	
CC	in mapping the naturally occurring genomic sequences and to create	
CC	knock-in humanised animals (pigs) or transgenic animals (mice or rats) to	
CC	model human diseases. GCRC is used to diagnose, prevent and treat	
CC	proliferative disorders (actinic keratosis, arteriosclerosis, cirrhosis,	
CC	hepatitis and cancer); cancer (breast, bladder, bone marrow, brain,	
CC	uterus cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma)	
CC	neurological disorders (epilepsy, stroke, Alzheimer's, Huntington's,	
CC	Parkinson's disease, multiple sclerosis, dementia and other central	
CC	nervous system disorders); cardiovascular disorders (angina pectoris,	
CC	hypertension, atherosclerosis, congestive heart failure);	
CC	gastrointestinal disorders (dysphagia, peptic oesophagitis, oesophageal	
CC	spasm, gastritis, gastric carcinoma, anorexia, nausea, abdominal angina,	
CC	pyrosis, pancreatitis, Crohn's disease, diarrhoea); autoimmune/	
CC	inflammatory disorders (acquired immunodeficiency syndrome (AIDS),	
CC	Addison's diseases, allergies, anaemia, asthma, diabetes mellitus, atopic	
CC	dermatitis, glomerulonephritis, Grave's disease, osteoarthritis,	
CC	psoriasis, rheumatoid arthritis, ulcerative colitis, bacterial, fungal,	
CC	parasitic, protozoal and helminthic infections) and metabolic disorders	
CC	(obesity, osteoporosis, viral infections).	
XX		
SQ	Sequence 3365 BP; 612 A; 1106 C; 1114 G; 533 T; 0 other;	
	Alignment Scores:	
	Prod. No.: 1..5e-269 Length: 3365	
	Score: 3218.00 Matches: 625	
	Percent Similarity: 100.00% Conservative: 0	
	Best Local Similarity: 100.00% Mismatches: 0	
	Query Match: 100.00% Indels: 0	
	DB: 22 Gaps: 0	
	US-10-029-347-2 (1-625) x AD089836 (1-3365)	
Qy	1 MetLeuA:aGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaSerGluLeuProA:a 20	
Dz	799 ATGCTGGCGCCAGCCGACGGACTGCTTTCA?CCTGGACGGCGGACGAGCTGCCGGCG 858	
Qy	2: LeuGlyGlyProGluA:laAProCyThrAspProPheGluAlaAlaSerGlyAlaArg 40	
Dz	859 CTTCGGGGGCCCCGAGGGCCGCCCCTTGCACAGACCCCTTCAGCGCGGACGCGCGCGCG 918	

```

QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuLeuValThrThr 60
DB 9-9 GTGCTAGCGGGCTGCTGAGCAAGCGCTGCTGCCACGCGCTCTGCTGGTACCAAG 978
QY 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 83
DB 979 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1038
QY 81 ArgGlyPheSerAspLysAspLysLysLysThrPheThrLysPhePheArgAspGluArg 103
DB 1039 CGCGCTTCTCCGACAAGGACAAGAAAGATATTTCTACAGATTTCTCCGGGATGATGAGG 1098
QY 101 ArgAlaGluArgAlaThrArgPheValLysGluLeuGlnThrLeuPheAlaLeuCysPhe 120
DB 1099 AGGGCGGAGCGCGCTACCGTTTCGTGAGGAGAGACGAGACGCTGTTCGGCGCTGCTTC 1158
QY 121 ValProPheValCysTrpPheValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
DB 1159 GTGCCCTTCGTGTGTGTGATATGTGTGTACCGCTGTGGCGCAGAGCTGGAGCTGGGTGG 1218
QY 141 AspLeuSerArgThrSerLysThrThrThrSerValThrLeuLeuPheLeuThrSerVal 160
DB 1219 GACCTGTCCGGCAGCTCCCAAGACCAACCAAGTACAGTGTACCTGCTTTTCATCACCAGCGTT 1278
QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAspLeuCys 180
DB 1279 CTGAGCTTCGGCTCCGCTAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTC 1338
QY 181 ArgLeuAlaArgGluGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
DB 1339 CGCCTGGCGCGCGAGGGCTCTCCGACGCGAGGGCGAGTTTTCGAGAGAGAACTGGAG 1398
QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
DB 1399 CAACCTGGAGACTTCGTGGCTCCAAAGTCAGACGCTGTTCTCAGCAAAAGAGAGCTCCCG 1458
QY 221 GlyValLeuGluThrGluValThrThrGlnPheLeuAspGlnSerPheGlnPheLeu 240
DB 1459 GCGGTGCTGGAGACAGAGGTCACTTACAGTTCATCGACAGAGCTTCAGAGAGTTCTCTC 1518
QY 241 AlaAlaLeuSerCysLeuLeuGluAspGlyValProAlaProAlaAlaGlyGlyVal 260
DB 1519 GCGGCACTGCTCTACTCTGCTGAGGACGCGGGGTGCCAGACACCGCGCTGGCGGCGTT 1578
QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
DB 1579 GGGACACTCTCTCGTGGGGACCGCCAGCGCGACAGCCACTTGGTGCTCACCACGCGCTTC 1638
QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHspheGlyCysMet 300
DB 1639 CTCTTCGGACTCTCTGCGCGCGAGCGGATGCGGACATCGAGCGCCACTTCGGGTGCATG 1698
QY 301 ValSerGluArgValLysGlnGlnAlaLeuArgTrpValGlnGlyGlnGlyCys 320
DB 1699 GTTCAGAGCGGTGTGAAGCAGAGAGCCCTCGCGTGGGTGCAGGGACAGGACAGCGGTGC 1758
QY 321 ProGluValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluPro 340
DB 1759 CCGGAGTGGACACAGAGTGAACGAGGGGGCCAAAGGGCTCGAGACACCCAGAGAGCCA 1818
QY 341 GluGluGluGluGluGlyGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
DB 1819 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1878
QY 361 GluThrGluGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
DB 1879 GAGACGAGGAGGAGCGCTTTGTGCCAAGCGCTGTGCCGCTTCCCGAGCTGGCGCTG 1938
QY 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCys 400
DB 1939 CAGCGAGTCCGCTTCTGCCGCAATGAGCGTGGCTGTCTGAGCTACTGCTGAGGTGCTGC 1998

```

```

QY 401 ProAlaGlyGlnAlaLeuArgLeuIleSerCysArgLeuValAlaAlaGlnGlnLysLys 420
DB 1999 CTGCTGAGCAGGCACTCGGCTGATCAGCTGCAGATTCGGTGTGCGCAGAGGAAG 2058
QY 421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlySerSerGlyGlyThr 440
DB 2059 AAGAAAGAGCCCTGGGAAAGCGCTCCAGGCGACCTGGGTGGCGGCACTCTCAAGGAC 2118
QY 441 ThrLysGlnLeuProAlaSerLeuHisProLeuPheGlnAlaMetThrAspProLeu 460
DB 2119 ACAAACAACCTGCCAGCTCCCTCTTTCATCCTCTTTCAGGCAATGACTGACCCACTG 2178
QY 461 CysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAsp 480
DB 2179 TGCCATCTGAGCAGCTCAGCTGTCCCACTGCCAACTCCCTGACCGGCTGCGGAGAC 2238
QY 481 LeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArg 500
DB 2239 CTTTCTGAGGCGCTGAGGCGAGCGCGCACTGACGAGGCTGGGCTCTCTCCACAACAGG 2298
QY 501 LeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgVal 520
DB 2299 CTCAGTGGGCGGAGCTGCTGATGCTGAGTGGGCGCTAGCTGGCGCAGTGCAGGGTG 2358
QY 521 GlnThrValArgValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGlyMet 540
DB 2359 CAGACGGTCAGGGTACAGCTGCCTGACCCCGCAGGAGGCTCCAGTACTGTTGGGTATG 2418
QY 541 LeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaPro 560
DB 2419 CTTGCGCAGAGCGCGCTGACCACTGATCTCAGCGCTGCCAATCTGCCCGCGCC 2478
QY 561 MetValThrThrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSer 580
DB 2479 ATGGTACCTACTCTGTGTGAGTCTGCGACGACCGAGGATGGGCTCGACACCTCAGT 2538
QY 581 LeuAlaSerValGluLeuSerGluGlnSerLeuGlnGluLeuGlnAlaValLysArgAla 600
DB 2539 CTGGCTCTGTGGAGCTGAGCGAGCAGTCACTACAGGAGCTCAGGCTGTGAAGAGAGCA 2598
QY 601 LysProAspLeuValIleThrHisProAlaLeuAspGlyHisProGlnProProLysGlu 620
DB 2599 AAGCGGATCTGTGTCATCACACACCGCGCTGGAGCGCCACCCACCACTCCCAAGGAA 2658
QY 621 LeuIleSerThrPhe 625
DB 2659 CTCATCTCGACCTTC 2673
RESULT 3
ABS63486
ID ABS63486 standard; cDNA; 2763 BP.
XX ABS63486;
AC ABS63486;
XX XX
XX -5-NOV-2002 (first entry)
XX XX
XX Human leucine-rich repeat small intestine I (HLRRS11), EST #1.
XX XX
XX Human leucine-rich repeat small intestine I; HLRRS11; asthma;
XX KW proliferative disorder; gastrointestinal disorder; renal disorder;
XX KW neural disorder; reproductive disorder; calcium regulation; apoptosis;
XX KW immune system; anaemia; human immune deficiency virus; HIV; cancer;
XX KW blood coagulation disorder; autoimmune disorder; allergic reaction;
XX KW inflammatory condition; cardiovascular disorder; ischaemia;
XX KW neurological disorder; infectious disease; cytokine production;
XX KW expressed sequence tag; EST; gene; ss.
XX OS Homo sapiens.
XX PN W0200261086-A2.
XX XX
XX 08-AUG-2002.
XX XX

```

```

PF 20-DEC-2001; 2001WO-US49739.
XX
XX
XX 22-DEC-2000; 2000US-257774P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Feder J, Ramanathan C, Mintier G;
XX WPI; 2002-619252/66.
XX
XX New isolated nucleic acid molecules encoding HLRRSII polypeptides, or
PT their fragments and homologues, useful for preventing, treating and
PT ameliorating medical conditions, e.g. proliferative, gastrointestinal,
PT or renal disorders
XX
XX Example 1; Page 221-222; 336pp; English.
XX
XX The invention relates to isolated nucleic acid molecules (I) encoding
CC human leucine-rich repeat small intestine - (HLRRSII) polypeptides.
CC The nucleic acid molecules and polypeptides are useful for preventing,
CC treating and ameliorating medical conditions, such as proliferative,
CC gastrointestinal, renal, neural, or reproductive disorders; or disorders
CC related to aberrant calcium regulation or apoptosis modulation, either
CC directly or indirectly. They are also useful for treating, preventing
CC and/or diagnosing diseases, disorders and/or conditions of immune system
CC by activating or inhibiting the proliferation, differentiation, or
CC mobilisation of immune cells; haematopoietic cells e.g. thrombocytopoietic,
CC anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
CC virus (HIV) infection, HTLV-ELV infection; blood coagulation disorders,
CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
CC myasthenia gravis; asthma or allergic reactions; inflammatory
CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
CC ischaemias, aneurysms; neurological disorders, e.g. Alzheimer's disease,
CC Huntington's chorea; infectious diseases, e.g. measles, mumps.
CC pneumonia, or viral, bacterial, and fungal infections. The HLRRSII
CC polypeptides are useful for modulating cytokine production, antigen
CC presentation, or other processes such as boosting immune responses.
CC ABS63485-ABS63504 represent HLRRSII coding sequences and PCR primers of
XX the invention.
XX
XX Sequence 2763 BP; 544 A; 890 C; 870 G; 459 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,7e-265 Length: 2763
XX Score: 3,70.00 Matches: 624
XX Percent Similarity: 96.00% Conservative: 0
XX Best Local Similarity: 96.00% Mismatches: 1
XX Query Match: 98.51% Indels: 26
XX DB: 24 Gaps: 1
XX
XX US-10-029-347-2 (1-625) x ABS63486 (1-2763)
XX
XX Qy 1 MetLeuAlaGlnProGluArgLeuLeuPheIleLeuAseGlyAlaAspGluLeuProAla 20
XX
XX Db 75 ATGCTGGCCACGCCGACGGCGTGCTCTTCATCTCGAGCGCGCGACGAGCTCGCGCG 34
XX
XX Qy 21 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAseGlyAlaArg 40
XX
XX Db 135 CTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 194
XX
XX Qy 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
XX
XX Db 195 GTGCTAGGGCGGGCTGCTGAGTAAAGCGCGTGTGTCACCGGGCCCTTCCTGTGGTACACG 254
XX
XX Qy 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlyVal 80
XX
XX Db 255 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 314
XX
XX Qy 81 ArgGlyPheSerAspLysAspLysLysLysTyrPheTyrLysPheArgAspGluArg 100
XX
XX Db 315 CGCGGGCTTCCTCGCAAGGACAAGGAAGAAGTATTTCCTACAGTTCCTTCGGGATGAGAG 374

```

Qy	101	ArgAlaGluArgAlaTyArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe	120
Db	375	AGGCGCGAGCGCGCTACCGCTTCGTGAAGCAGACGAGCGGTGTTCGCGCTGTGCTTC	434
Qy	122	ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluArg	140
Db	435	GTGCGCTTCGTGTGCTGGATCGTGTGCACCGTGTGCGCCAGCAGCTGGAGCTCGGTTCGG	494
Qy	141	AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal	169
Db	495	GACCTGTGCGCGACTCCAGACCAACCGTCACTGTACCTGCTTTTCAACACAGCGTT	554
Qy	161	LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys	180
Db	555	CTGAGCTCGGCTCGGTAGCGAGCGGCGCCCGGTTCAGGCGCACTGCGCAATCTGTGC	614
Qy	181	ArgLeuAlaArgGluGlyValLeuGlyArgAlaGlnPheAlaGluTysGluLeuGlu	200
Db	615	CGCTGCGCGCGAGGCGCTCTCGAGCGAGGCGCGAGTTTCCGAGAAAGGAACCTGGAG	674
Qy	201	GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro	220
Db	675	CAACTGGAGCTTCGTGGCTCCAAAGTCGACACCGCTGTTCTCAGACAAAAGGAGCTGCCG	734
Qy	221	GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu	240
Db	735	GGCGTGTGGAGACAGAGGTCACTACCACTTCATCGACCAGAGCTTCCAGGAGTTCCCTC	794
Qy	241	AlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyVal	260
Db	795	GGCGCACTGTTCCTGTGGAGGACGGCGGGGTGCCAGGACCGCGGTGCGCGCGCTT	854
Qy	261	GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe	280
Db	855	GGGACACTCTCGTGGGAGCGCCAGCCGACAGCCACTTGGTGTCTCACACGCGCTTC	914
Qy	281	LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet	300
Db	915	CTCTTCGGACTGTGAGCGCGGAGCGGATCGCGCACTCAGCGCCACTTCGGCTGCATG	974
Qy	301	ValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlnGlyGlnGlyCys	320
Db	975	GTTCACAGCGCTGTGAAGCAGAGGAGCGCTCGCGTGGGTGCAGGCGACAGGCGCTGC	1034
Qy	321	ProGlyValAlaProGluValThrGluGlyValLysGlyLeuGluAspThrGluGluPro	340
Db	1035	CCCGAGTGCACACAGAGTGTACCGAGGGGGCCAAAGGGCTCAGGAGACCGAAGAGCCA	1094
Qy	341	GluGluGluGluGluGlyGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr	360
Db	1095	GAGGAGGAGGAGGAGGAGGAGGAGCCCACTACCCACTGGAGTTGCTGTACTGCCGTAC	1154
Qy	361	GluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu	380
Db	1155	GAGACGCAGAGAGACGGCTTGTGCGCCCAAGCCCTGTGCGCGTTCCCGGAGCTGGCGCTG	1214
Qy	381	GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys	400
Db	1215	CAGCGAGTGGCTTCGCGCCATGGAGCTGGCTGTCTTGAGCTACTGCGTGAAGTGTCTGC	1274
Qy	401	ProAlaGlyGlnAlaLeuArgLeuIleSerCysArgLeuValAlaAlaGlnGluLysLys	420
Db	1275	CCTGCTGGACAGGCACTGCGCGCTGATCAGCTGCAGATTGGTTGCTCGCAGGAGAGAAG	1334
Qy	421	LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlySerSerGlnGlyThr	440
Db	1335	AAGAAGAGCCTGGGGAAGCGGCTTCAGGCGCAGCCTGGGTGGCGGCGAGTCTCAAGGCACC	1394
Qy	441	ThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeu	460
Db	1395	ACAAAACACTGCCGCGCTCCCTCTTCATCCACTCTTTCAGGCGAATGACTGACCCACTG	1454
Qy	461	CysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAsp	480

```

Db      1455  TGGCATTCTGAGCAGCCTCAAGCTGTCCCTCACTGCAAACTCCCTCACCAGGCTCTGCCAGAGAC 1514
Qy      481  LeuSerGluAlaLeuAgaAaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArg 500
Dt      1515  CTTTCTGAGGCCCTGAGGGVAGCCCGCACTGAGGAGCTGGCCCTCTCCACACAGG 1574
Qy      501  LeuSerGluAlaGlyLeuA-gMetLeuSerGluGlyLeuAlaTrpProGlnCysAsVal 520
Db      1575  CTCAGTGAAGCGGAGCTGCTATGCTGAGTGAGGGCCCTAGCTGGCCGCGCAGTGCAGGGTG 1634
Qy      521  GlnThrVal----- 523
Db      1635  CAGACGGT- CAGGTGAGGCTTGGCTGGGAGGGACCGTGGGATGCCCCGCCACCCAGC 1693
Qy      524  -----ArgValGlnLeuProAspProGlnArgGlyLeuGln 535
Db      1694  AGCTCCTGAGGTGGCCCTCCACAGGGTACAGCTGCCTCAGCCCGCAGGAGGGCTCCAG 1753
Qy      536  TyrLeuValGlyMetLeuA-gGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCys 555
Db      1754  TACCTGGTGGTATGCTTGGCAGAGCCCGCCCTGACCACTGATCTTCAGCGGCTGC 1813
Qy      556  GlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGly 575
Db      1814  CAACCTGCCCGCCCATGGTGACCTACTGTGTGCAGTCTGCGAGCACCGAGGATCGGC 1873
Qy      576  LeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnGluLeuGln 595
Db      1874  CTGAGACCTCTCAGTCTGGGCTCTGTGGAGCTGAGCGAGCAGTCACTACAGGAGCTTCAG 1933
Qy      596  AlaValLysArgAlaLysProAspLeuValLeuThrHisProAlaLeuAspGlyHisPro 615
Db      1934  GCTGTGAAGAGAGCAAGCCGGATCTGCTATCAACACACCCAGCGCTGGACGCCACCCCA 1993
Qy      616  GlnProProLysGluLeuLeuSerThrPhe 625
Db      1994  CAACCTCCCAAGGAACCTCATCTGCACCTTC 2023

RESULT 4
ABX97042
ID      ABX97042 standard; cDNA; 2109 BP.
AC      ABX97042;
CC      20-MAR-2003 (first entry)
DE      Human NOV18a cDNA.
KW      NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW      hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW      human; gene; ss.
OS      Homo sapiens.
PN      WC200272757-A2.
PD      19-SEP-2002.
PF      C8-MAR-2002; 2002WC-US06908.
XX      PR
PR      08-MAR-2001; 2001US-274101P.
PR      08-MAR-2001; 2001US-274194P.
PR      08-MAR-2001; 2001US-274281P.
PR      08-MAR-2001; 2001US-274322P.
PR      09-MAR-2001; 2001US-274849P.
PR      12-MAR-2001; 2001US-275235P.
PR      13-MAR-2001; 2001US-275578P.
PR      13-MAR-2001; 2001US-275579P.
PR      13-MAR-2001; 2001US-275601P.
PR      14-MAR-2001; 2001US-276000P.
PR      16-MAR-2001; 2001US-276769P.
PR      19-MAR-2001; 2001US-276994P.
PR      20-MAR-2001; 2001US-277239P.
PR      20-MAR-2001; 2001US-277321P.
PR      20-MAR-2001; 2001US-277327P.
PR      21-MAR-2001; 2001US-277791P.
PR      22-MAR-2001; 2001US-277833P.
PR      23-MAR-2001; 2001US-278152P.
PR      26-MAR-2001; 2001US-278894P.
PR      27-MAR-2001; 2001US-278999P.
PR      27-MAR-2001; 2001US-279036P.
PR      28-MAR-2001; 2001US-279344P.
PR      30-MAR-2001; 2001US-277338P.
PR      30-MAR-2001; 2001US-279995P.
PR      30-MAR-2001; 2001US-280233P.
PR      30-MAR-2001; 2001US-280802P.
PR      02-APR-2001; 2001US-280822P.
PR      02-APR-2001; 2001US-280900P.
PR      04-APR-2001; 2001US-281194P.
PR      13-APR-2001; 2001US-283675P.
PR      30-APR-2001; 2001US-287424P.
PR      02-MAY-2001; 2001US-288066P.
PR      03-MAY-2001; 2001US-288342P.
PR      03-MAY-2001; 2001US-288528P.
PR      15-MAY-2001; 2001US-291190P.
PR      16-MAY-2001; 2001US-291099P.
PR      16-MAY-2001; 2001US-291240P.
PR      30-MAY-2001; 2001US-294485P.
PR      31-MAY-2001; 2001US-294889P.
PR      31-MAY-2001; 2001US-294899P.
PR      18-JUN-2001; 2001US-299027P.
PR      19-JUN-2001; 2001US-299303P.
PR      19-JUN-2001; 2001US-299310P.
PR      10-JUL-2001; 2001US-304354P.
PR      31-JUL-2001; 2001US-309198P.
PR      16-AUG-2001; 2001US-312903P.
PR      10-SEP-2001; 2001US-318462P.
PR      12-SEP-2001; 2001US-318770P.
PR      27-SEP-2001; 2001US-325430P.
PR      27-SEP-2001; 2001US-325681P.
PR      18-OCT-2001; 2001US-330380P.
PR      31-OCT-2001; 2001US-335301P.
PR      14-NOV-2001; 2001US-332172P.
PR      14-NOV-2001; 2001US-332272P.
PR      14-NOV-2001; 2001US-332722P.
PR      14-NOV-2001; 2001US-333184P.
PR      14-NOV-2001; 2001US-333272P.
PR      21-NOV-2001; 2001US-332694P.
PR      03-DEC-2001; 2001US-337426P.
PR      03-DEC-2001; 2001US-338092P.
PR      04-DEC-2001; 2001US-337185P.
PR      03-JAN-2002; 2002US-345705P.
PR      07-MAR-2002; 2002US-0092900.
XX      (CURA-) CURAGEN CORP.
XX      PA
XX      Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
XX      Zerbussen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
XX      Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
XX      Fernandes ER, Casmar S, Malyankar UM, Gerlach V, Liu Y;
XX      Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H;
XX      Alsobrook JP, Lepley DM, Rieger DK;
XX      MPI; 2002-723332/78.
XX      P-PSDB; ABU65075.
XX      DR
XX      NOVX polypeptides and polynucleotides, useful for preventing or
XX      treating a disorder associated with aberrant NOVX expression or
XX      activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
XX      bronchial asthma.
XX      PS
XX      Claim 13; Page 149; 1103pp; English.
XX      This invention describes novel human NOVX polypeptides which have
XX      cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and
XX      CC

```


CC hypotensive activity. Pharmaceutical compositions comprising the NOVX
 CC proteins or nucleic acid molecules or NOVX antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOVX
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABX97008-ABX97185 are cDNA
 CC fragments amplified and isolated by the PCR primers and probes
 CC represented in ABX13460-ABX13462 and ABX97186-ABX97593. ABX97008-ABX97185
 CC encode the NOVX proteins described in ABU65041-ABU65218.

XX Sequence 2109 BP; 371 A; 673 C; 717 G; 348 T; 0 other;

Alignment Scores:

Pred. No.: 2,13e-263 Length: 2,109
 Score: 3,144.00 Matches: 6,15
 Percent Similarity: 98.40% Conservative: 0
 Best Local Similarity: 98.40% Mismatches: 4
 Query Match: 97.70% Indels: 6
 DB: 24 Gaps: 1

US-10-029-347-2 (1-625) x ABX97042 (1-2109)

QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
 DB 197 ATGCTGGCCAGCCAGCGGCTGCTTCTATCTGGACGGCGCGGACGAGCTGCCGGCG 256
 QY 21 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
 DB 257 CTGGGGGCGCCGAGCGCCCTGGCAGACGCCCTTCGAGGGCGGAGGGCGCGCGG 316
 QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrTr 60
 DB 317 GTGCTAGCGGGCTGCTGAGTAAGCGCTGCTGCCACCGGCCCTCTGCTGGTGCACCG 376
 QY 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
 DB 377 CGCGCGCGCGCCCGGAGGCTGCAGGGCGCGCTGTGTTCGCCGAGTGGCGCGAGGTG 436
 QY 81 ArgGlyPheSerAspLysAspLysLysLysLysPheThrLysPhePheAspGluArg 100
 DB 437 CGCGCTTCTCCGACAGGACCAAGAGAGTATTTCTACAAGTTCTTCGGGATGAGAGG 496
 QY 101 ArgAlaGluArgAlaArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
 DB 497 AGCGCGGAGCGCCCTACCGCTCTGTAAGAGAGACGAGAGCTGTTCCGCTGTGCTTC 556
 QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
 DB 557 GTGCCCTTCTGTTGCTGGATCGTGTGCACCGCTGCTGCCGACGAGCTGGAGCTCGGTCGG 616
 QY 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
 DB 617 GACCTGTCCGCGACGCTCAAGACCAACACGTCAGTGTACCTGCTTTTCATCACCAGCGTT 676
 QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
 DB 677 CTGAGCTTGGCTCCGCTAGCCAGCGGCCCGCTTGCAGGGGCGACCTCGGCAATCTGTCG 736
 QY 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
 DB 737 GCGCTGGCGCGGAGGGGCTCTCCGAGCGAGGGCGCAGTTTTCGCGAAGAGAACTGGAG 796
 QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
 DB 797 CAACCTGGAGCTTCTGGCTCCAAAGTGACAGCGCTGTTTCTCAGCAAAAGAGAGCTGCCG 856
 QY 221 GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
 DB 857 GCGCTGCTGGAGACAGAGGTCACCTACCAATTCTCGACACGAGGCTTCAGGAGTCTCTC 916
 QY 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
 DB 917 GCGGCACTGTCTCTGCTGGAGGACGCGCGGGGTGCCAGGACCGCGGCTGGCGGGCTT 976

QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
 DB 977 GGACACTCTCTGGTGGGAGCGCCAGCGCAGCAGCTTGTGTCTACCCAGCGCTTC 1036
 QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
 DB 1037 CTCTTCGGACTCTGAGCGCGGAGCGGATGCGCGACATCGAGCGCCACTTCGCGTCATG 1096
 QY 301 ValSerGluArgValLysGlnGluAlaLeuArgTrpValGlnGlyGlnGlyGlnGlyCys 320
 DB 1097 GTTTCAGAGCGTGTGAAGCAGGAGGCGCTGCGGTGGGTGCGAGGACAGGAGCGCTGC 1156
 QY 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
 DB 1157 CCGGAGTGGCACCAGAGGTCAGGAGGGGCGCAAGGGCTCGAGGACACCCAGAGAGCCA 1216
 QY 341 GluGluGluGluGluGlyGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
 DB 1217 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1276
 QY 361 GluThrGlnGlnAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
 DB 1277 GAGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1336
 QY 381 GluArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
 DB 1337 CAGCGAGTGGCTTCTGCCCATGAGCTGGCTGTTCTGAGCTACTGCTGAGTGGTGGCTGC 1396
 QY 401 ProAlaGlyGlnAlaLeuArgLeuIleSerCysArgLeuValAlaAlaGlnGluLysLys 420
 DB 1397 CTTGCTGCACAGGCACTGGCGGTGATCAGTGCAGATGTTGTTCTGCGCAGGAGAAGAAG 1456
 QY 421 LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyCysSerSerGlnGlyThr 440
 DB 1457 AAGAAAGAGCTTGGGGAAGCGGCTCCAGGCGCAGGCTG-----GGCACC 1498
 QY 441 ThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeu 460
 DB 1499 ACANAACAATGCCAGCTCCCTTCTTCATCTCACTCTTTCAGGCAATGACTGACCCACTG 1558
 QY 461 CysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAsp 480
 DB 1559 TGCCATCTGAGCAGCTCAGCTGTCCCACTGCAAACTCCCTCAGCGGCTCTCCGAGAC 1618
 QY 481 LeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArg 500
 DB 1619 CTTTCTGAGGCGCTGAGGGCAGCGCCCGCACTGACGAGCTGGGCGCTCTCCACAACAGG 1678
 QY 501 LeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgVal 520
 DB 1679 CTGAGTGAGCAGAGCTGCGTATGCTGAGTGGGCTGAGCTGGCGCGAGTGCAGGGTG 1738
 QY 521 GlnThrValArgValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGlyMet 540
 DB 1739 CAGCGCTCAGGCTACAGCTGCTGACCCCGCAGCGGGCTCCAGTACCTGGTGGGTATG 1798
 QY 541 LeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaPro 560
 DB 1799 CTTCCGCGAGAGCGCTCGCTGACCCCTGAGCAGCTGAGTCTCAGCGGCTGCCAACCGCCC 1858
 QY 561 MetValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSer 580
 DB 1859 ATGCTGACCTACTCTGTGTGAGCTCTTCAGCAGCAGGAGTGGCGCTCGAGACCTCAGT 1918
 QY 581 LeuAlaSerValGluLeuSerGluGlnSerLeuGlnGluLeuAlaValLysArgAla 600
 DB 1919 CTGGCTCTCTGGAGCTGAGCGAGAGCTCACTACAGAGCTTCAGGCTGTGAAGAGAGCA 1978
 QY 601 LysProAspLeuValIleThrHisProAlaLeuAspGlyHisProGlnProLysGlu 620
 DB 1979 AAGCCGATCTGTGTCTCATCACACCCAGCGCTGGAGCGGCACCCCAACCTCCCAAGGAA 2038

QY 621 LeuIleSerThrPhe 625
 DB 2039 CTCATCTCGACCTTC 2053
 RESULT 5
 NAL47132
 ID AAL47132 standard; DNA; 5556 BP.
 XX
 AC
 XX
 DT 20-AUG-2002 (first entry)
 XX
 DE Pyrin domain containing protein NALP6/PY9 coding sequence.
 XX
 KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
 KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis; gene; ds.
 XX
 CS Unidentified.
 XX
 FN WO200240568-A2.
 XX
 PD 23-MAY-2002.
 XX
 PF 30-OCT-2001; 2001WO-EP:2545.
 XX
 PR 15-NOV-2000; 2000DE-1056687.
 PR 30-NOV-2000; 2000DE-1059595.
 XX
 PA (APOT-) APOTECH RES & DEV LTD.
 XX
 PI Tschopp C, Martinen F;
 XX
 DR WPI: 2002-427093/45.
 DR P-PSDB; RAO17860.
 XX
 PT New DNA encoding protein: with pyrin domain, useful for treating
 PT diseases involving impaired signal transduction, particularly
 PT inflammation, also proteins and antibodies
 XX
 PS Claim 5; Fig 1; 116pp; German.
 XX
 CC The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis.
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a coding sequence of
 CC the invention.
 XX
 SQ Sequence 5556 BP; 914 A; 1856 C; 1879 G; 957 T; 0 other;
 Alignment Scores:
 Pred. No.: 5,01e-24C Length: 5556
 Score: 2882.00 Matches: 551
 Percent Similarity: 82.82% Conservative: 7
 Best Local Similarity: 81.83% Mismatches: 32
 Query Match: 89.56% Indels: 90
 DB: 24 Gaps: 5
 US-10-029-347-2 (1-625) x AAL47132 (1-5556)
 QY 1 MetLeuAlaGlnProGlnArgLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
 DB 799 ATGCTGCCACCGCCAGCGGCTCTCTTCATCTCTGACCGCGGACGAGCTGCCGCGG 858
 QY 21 LeuGlyGlyProGluAlaAlaProCysThrAspPheGluAlaAlaSerGlyAlaArg 40

DB 859 CTGGGGGGCCCCAGAGCGCGCCCTTCACAGACCCCTTCAGCGCGAGCGCGCGG 918
 QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
 DB 919 GTGCTAGCGGGCTGCTGAGCAGAGGCGCTCTGCCACGSCCTCTCTCTCTGTCACACG 978
 QY 61 ArgAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
 DB 979 CGCGCGCGCGCGCGCGGCTGACGAGCGCGCTGTGTTCCTCCCGCAGTGTGCGCGAGGTG 1038
 QY 81 ArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLys 100
 DB 1039 CGCGGCTTCTCCGACAAAGGACAAAGAGAGTATTTCTACAAAGTCTTCGCGGATGAGAGG 1098
 QY 101 ArgAlaGluArgAlaTyrArgPheValLysGluAsnGlnThrLeuPheAlaLeuCysPhe 120
 DB 1099 AGGCGCGAGCGCGCTACCGCTTCGTGAAGAGACGAGACGCTGTTCGCGCTGTCTTC 1158
 QY 121 ValProPheValCysThrPheValCysThrValLeuArgGlnGlnGlnGlnGlnGlyArg 140
 DB 1159 GTGCCCTTCGTGTCTGTGATCGTGTGCACCGTGTGCGCGCAGAGCTGGAGCTCGTCCG 1218
 QY 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
 DB 1219 GACCTGTCCGCGCCTCCAGACCCACCGCTCAGTGTACTGTCTTTCATCACCAGCGTT 1278
 QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
 DB 1279 CTGAGCTCGGCTCGGTAGCGCGCGCGCGCTTCGAGCGCGACCTTCGCAATCTGTGC 1338
 QY 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
 DB 1339 GGCCTGGCGCGAGGCGCTCCGAGCGCGCGCGCGCTTCGCGAGAGGAACTCGGAG 1398
 QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
 DB 1399 CACTGGAGCTTCGTGCTCCAAAGTGCACAGCTGTCTCAGCAAAAGAGAGTGCCTG 1458
 QY 221 GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
 DB 1459 GGCTGTCTGGAGCAGAGGTCACTACAGTTTATCCACAGAGCTTCAGAGGATTCCTC 1518
 QY 241 AlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
 DB 1519 CGCGCATGTCTTACCTGTCTGGAGCGCGCGGTGCCAGAGCCGCGCTGGCGCGCTT 1578
 QY 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
 DB 1579 GGGACATCTCTGTGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1638
 QY 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
 DB 1639 CTCCTCGACTGTCTGAGCGCGGAGCGGATCGCGGACATCGAGCGCCACTTCGCGCTG 1698
 QY 301 ValSerGluArgValLysGlnGluAlaLeuArgTyrValGlnGlyGlnGlyGlnGlyCys 320
 DB 1699 GTTTCAGAGCGTGTGAAGCAGGAGCGCGCTGCGGTGGGTGCAGGACAGGAGCGGCTGC 1758
 QY 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340
 DB 1759 CCCGAGTGGCACCAGAGGTGACCGAGGGGCGCCAAAGGGCTCGAGGACACCGAAGGCCA 1818
 QY 341 GluGluGluGluGluGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
 DB 1819 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1878
 QY 361 GluThrGlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
 DB 1879 GAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1938
 QY 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400

```
Db 1939 CAGGAGTGGCTTCTCCCGATGGAGCTGGTGTCTTGAGCTACTGCGTGAGGTGCTGC 1998
Qy ProAlaGlyGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaGlnGulysLys 420
Db 1999 CCGTGGACAGGCACTGGCGCTGATCAGCTGCAGATTGGTCTGGCAGGAGAGAG 2058
Qy LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlySerSerGlnGlyThr 440
Db 2059 AAGAAGAGCCTGGGAAGCGCTCCAGGCCAGCCTGGTGGCGAGCTGGCTGGGGACC 2118
Qy 440 ----- 440
Db 2119 CAACCTGCTCAGAAGTACCTTTTCGACACACCTGCTGTGACATCTGCCCCACACCTCCA 2178
Qy 440 ----- 440
Db 2179 CCAGACCTCGGCTCCTCCAGGGCAAGGCTTTTCCAGAGACTTCTTTGAATATAGCTCCA 2238
Qy 441 ThrLysGlnLeuProAlaSerLeu----- 449
Db 2239 ATTACGCCCTGCCCGAGGGCTTGGCATCTTTCAGAGGATGAATGTCAGCGTGTGGCA 2298
Qy 450 -----HisProLeuPheGlnAlaMetThrAspProLeuOysHisLeuSer 454
Db 2299 GGGCTGGGCTGGGACCCAAAGACCATGCAATGACTGACCATCTGTCACATCTGAGC 2358
Qy 465 SerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerGluAla 434
Db 2359 AGCCTCAGCTGTCCCACTGCAAACTCCCTGACCGCTCCCGAGACCTTTCTGAGGCC 2418
Qy 485 LeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAla 504
Db 2419 CTGAGGGCAGCCCGCACTGACGAGCTGGGCTCTCTCCACAAACAGGCTCAGTGAGGCA 2478
Qy 505 GlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlcCysArgValGlnThrValArg 524
Db 2479 GGAATCGTATGCTGAGTAGGGCCCTAGCCTGGCCGACGTGCAGCGTGCAGAGCGG-CAGG 2538
Qy 525 ValGlnLeuProAspProGlnArgGlyLeuGlnTyLeuValGlyValLeuArgGlnSer 544
Db 2539 GTACAGCTGCTGACCCCAAGCAGCGGCTCCAGTACTGCTGGTATGCTTCGGCAGAGC 2598
Qy 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTy 564
Db 2599 CTTGCCCTGACCACTGATCTCAGCGCTGCACTGCGCCGCCCCCACTGGTGGACCTAC 2658
Qy 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAla----- 582
Db 2659 CTGTGTGCACTCTGCAGCACCAGGATGGGCTGCAGACCTCAGCCTCTGCTGCTTCT 2718
Qy 583 -----SerValGluLeuSerGlu----- 588
Db 2719 TCTGACCGACCCGAGTCTTCTCGGAGCGTGTGAGAACCCGGCGCGGCTGGGG 2778
Qy 589 -----GlnSerLeuGlnGluLeuGlnAlaValLysArgAlaLysPro 602
Db 2779 CTGGAGTCTCGTGGCTCGGAGCGCCCGCGACCTCGGGCCACAGCGAGGAGGACC 2838
Qy 603 AspLeuValIleThr-----HisProAla 610
Db 2839 CAGTGGAGCGCGCGGGCGCGGGGAGGAGGAGCGGGAGGAAACCCCGGCGC 2898
Qy 611 LeuAspGlyHisProGlnProProLysGlu 620
Db 2899 CGCGGGCCCAACCAACAGCGCGCCCGGGGAC 2928
```

RESULT 6

AAS00033

ID AAS00033 standard; cDNA; 6461 BP.

AC AAS00033;

XX AAS00033;

DT 10-MAY-2001 (first entry)

```
XX Human ATLAS-2-encoding cDNA.
DE Human: Activated T-lymphocyte associated sequence 2; ATLAS-2; antibody;
KW Cytokine receptor; autoimmune disorder; immune disorder; cancer;
KW T-lymphocyte-associated disorder; cell-proliferation disorder; tumour;
KW cell differentiation disorder; immune deficiency disorder; malignancy;
KW viral infection; bacterial infection; fungal infection; metabolism;
KW chromosome 11p15.5; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..5556
FT /tag= a
FT /product= "ATLAS-2"
XX
FN WC200114564-A2.
XX
PD 01-MAR-2001.
XX
PF 18-AUG-2000; 2000MO-US22699.
XX
PR 20-AUG-1999; 99US-0150105.
PR 28-APR-2000; 2000US-0560101.
PR 28-APR-2000; 2000US-0560365.
PR 28-APR-2000; 2000US-0560948.
PR 28-APR-2000; 2000US-0561533.
XX
PA (CURA-) CURAGEN CORP.
PA (BIO-) BIOGEN INC.
PI Feyman JA, Green CD, Hsu A, Browning JA, Carulli J;
XX
DR WPI: 2001-218453/22.
DR P-PSDB; AAU000023.
XX
PT New isolated activated T lymphocyte associated sequences for treating
PT or preventing immune system associated disorders such as autoimmune
PT disorder, immune disorder, and T-lymphocyte-associated disorder
XX
PS Claim 6; Fig 2; 114pp; English.
XX
CC The sequence encodes human activated T-lymphocyte associated sequence 2,
CC ATLAS-2. ATLAS-2 is related by homology to cytokine receptors and its
CC gene is located on chromosome 11p15.5. ATLAS proteins, polynucleotides
CC and antibodies are useful for treating/preventing conditions associated
CC with an autoimmune disorder, immune disorder, T-lymphocyte-associated
CC disorder, cell-proliferation disorder, cell differentiation disorder,
CC and immune deficiency disorder and for screening for a modulator of
CC activity or of latency or predisposition to an immune disorder.
CC ATLAS proteins, polynucleotides and antibodies are useful in therapeutic
CC or prophylactic treatment of diseases associated with cell-proliferation
CC (e.g. cancers, malignancies and tumours). The polynucleotides are useful
CC in gene therapy, to detect ATLAS mRNA or a genetic lesion in an ATLAS
CC gene, to modulate ATLAS activity, to screen drugs or compounds that
CC modulate ATLAS activity or expression and to treat disorders
CC characterised by insufficient or excessive production of ATLAS protein or
CC production of ATLAS protein forms that have decreased or aberrant
CC activity compared to ATLAS wild type protein and in tissue typing to
CC identify individuals. The antibodies are useful for localisation/
CC quantitation, isolation and detection of ATLAS and to monitor protein
CC levels in tissue. ATLAS is useful for treating/preventing infection by
CC bacteria, viruses and fungi, affecting bodily characteristics, e.g.
CC biorhythms, fertility or metabolism, affecting behavioural
CC characteristics, and for providing analgesic effects. A host cell
CC containing the polynucleotide is useful to produce non-human transgenic
CC animals.
XX
SQ Sequence 6461 BP; 1079 A; 2146 C; 2150 G; 1086 T; 0 other;
```

Alignment Scores:

Pred. No.: 6.19e-240 Length: 6461

Score: 2862.00 Matches: 58:
 Percent Similarity: 82.82% Conservative: 7
 Best Local Similarity: 81.83% Mismatches: 32
 Query Match: 89.56% Indels: 90
 DB: 22 Gaps: 5

US-10-029-347-2 (-625) x AAS00033 (1-6461)

Qy : Met:LeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeu:ProAla 20
 Db 799 ATGCTGGCCAGCGCGAGCGGCTGTCTTCTCATCTTGAAGCGCGAGCGAGCTGCGGCG 858
 Qy 21 LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAlaArg 40
 Db 859 CTGGGGGGGCGCGAGCGCGCTGCTGACAGAGCCCTTCAGAGCGGCGAGCGCGCGCG 918
 Qy 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuLeuValThrThr 60
 Db 919 GTGCTAGGCGGGCTGTGAGCAAGCGCTGTGCTGCCACCGGCTCTCTGCTGTGACACG 978
 Qy 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
 Db 979 CGCGCGCGCGCGCGAGGCTGACGGGCGCGCTGTGTTCGCCGAGTGCSCCGAGGTG 1038
 Qy 81 ArgGlyPheSerAspLysAspLysLysLysLysLysLysLysLysLysLysLysLysLys 100
 Db 1039 CGCGGCTTCTCCGACAAAGGACAAAGAAAGATATTTCTCAAGTGTCTTCGGGATGAGAGG 1098
 Qy 101 ArgAlaGluArgAlaTyArgPheValLysGluLysGlnGluThrLeuPheAlaLeuCysPhe 120
 Db 1099 AGSGCGGAGCGCGCTACCGCTTCGTGAGAGGAGACGACGCTGTTCGGCTGTGCTTC 1158
 Qy 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
 Db 1159 GTCCCTTCTGTGTGTGGAACGTGTGCACCGTGTGCGCCAGCGCTGAGGCTCGGTGCG 1218
 Qy 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
 Db 1219 GACTGTGCGGCGAGCTTCAGACACACACGCTCAGTGTACTGCTTTTATCACCAGCGT 1278
 Qy 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
 Db 1279 CTGAGTCTGGCTTCGGTAGCGAGCGGCGCGGCTTGCAGGGGAGCTGCGCAATCTGTGC 1338
 Qy 181 ArgLeuAlaArgGluGlyValLeuGlyValArgArgAlaGlnPheAlaGluLysGluGlu 200
 Db 1339 CGCTTGGCGCGGAGGCGTCTCGAGCGGAGGCGCGCTTGCAGAGAGGAACTGAG 1398
 Qy 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
 Db 1399 CAACTGGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTTTCTCAGCAAAAAGGAGCTGCG 1458
 Qy 221 GlyValLeuGluThrGluValThrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
 Db 1459 GCGGTGCTGGAGACAGAGTCACTACCAAGTTCATCGACAGAGCTTCCAGAGTTCCTC 1518
 Qy 241 AlaAlaLeuSerTyLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyVal 260
 Db 1519 CGCGCACTGCTTACTGCTGGAGAGAGCGGCGGTGCGGAGACCGCGGCTGGCGGCT 1578
 Qy 261 GlyThrLeuLeuArgGlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPhe 280
 Db 1579 GGGACACTCTGTGTGGGAGCGCGGAGCGGAGCGGAGCGGAGCTGTGTGTCTACCAAGGCTTC 1638
 Qy 281 LeuPheGlyLeuLeuSerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMet 300
 Db 1639 CTCTTCGGACTGTGTAGCGGAGCGGATGCGGACATCGAGCGCACTTCGCGCTGCATG 1698
 Qy 301 ValSerGluArgValLysGlnAlaLeuAlaArgTyrValGlnGlyGlnGlyGlnGlyCys 320
 Db 1699 GTTTCAGAGCGGTGTGAAGCAGAGAGCGCTGCGGTGAGTGCAGGAGCAGGAGCGGCTGC 1758
 Qy 321 ProGlyValAlaProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluPro 340

Db 1759 CCGGAGTGGCACACAGAGTGCAGAGGGGGCAAGGGCTCGAGACACCCAAAGAGCCA 1818
 Qy 341 GluGluGluGluGluGlyGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyr 360
 Db 1819 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1878
 Qy 361 GluThrGlnGluAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeu 380
 Db 1879 GAGACGACGAGGAGCGCTTGTGTGCCCAAGCGCTGTGCCGCTTCCCGAGTGGCGCTG 1938
 Qy 381 GlnArgValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCys 400
 Db 1939 CAGCGAGTGGCTTCTGCCGATGAGCGTGGCTTCTGAGTACTGCGTGGAGTGTCTGC 1998
 Qy 401 ProAlaGlyGlnAlaLeuArgLeuIleSerCysArgLeuValAlaAlaGlnGluLysLys 420
 Db 1999 CTTGTGGACAGGCACTGCGGCTGATCAGCTGCAGATTGTTGCTGCGCAGGAGAAGAG 2058
 Qy 421 LysCysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlySerSerGlnGlyThr 440
 Db 2059 AAGAAGAGCTTGGGGAAGCGGCTCCAGGCGACGCTGGTGGCGGAGCTGGCTGGGACCC 2118
 Qy 440 ----- 440
 Db 2119 CAACTGGCTCCAGAGTACCTTTCCAGCACCCCTGCTGTGACATCTGCCCCACACCTCCA 2178
 Qy 440 ----- 440
 Db 2179 CCAGACCTCGGCTCTCCAGGCAAGGCTTTTCCAGAGTTCCTTTGAATATATAGCTCCA 2238
 Qy 441 ThrLysGlnLeuProAlaSerLeuLeu ----- 449
 Db 2239 ATTACGCCCCCTGCCAGGCGCTTGGCTCTGTTGAGAGGATGAATGTCAGGTTGGCA 2298
 Qy 450 -----HisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSer 464
 Db 2299 GGGGCTGGGCTGGGACCCAAAGACCCATGCAATGACTGACCCACTGTGCCATCTGAGC 2358
 Qy 465 SerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerGluAla 484
 Db 2359 AGGCTCACGCTGTCCCACTGCCAACTCCCTGACGGCTCTGCGGAGACCTTTCTGAGGCC 2418
 Qy 485 LeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAla 504
 Db 2419 CTAGGGGACGCGCGCACTGACGAGCTGGGCTCTCTCCACACAGGCTCAGTGAAGCA 2478
 Qy 505 GlyLeuArgMetLeuSerGluGlyLeuAlaTyrProGlnCysArgValGlnThrValArg 524
 Db 2479 GGACTGCTATGCTGAGTGGGCTTAGCTGGCGCAGTGCAGGCTGCAGAGCTCAGG 2538
 Qy 525 ValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGlyMetLeuArgGlnSer 544
 Db 2539 GTACAGTGTCTGACCCCGCAGCGAGGCTCCAGTACTGTTGGGTATGCTTCGGCAGAGC 2598
 Qy 545 ProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyr 564
 Db 2599 CTTGCCCTGACCACTGGATCTCAGCGCTGCCAACTGCCCGCCCGCATGGTGACCTAC 2658
 Qy 565 LeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSerLeuAla ----- 582
 Db 2659 CTGTGTGCTGCTGAGCAGCACCGAGGATGGGCTGCGACCTCAGACCTCTGCTCTCT 2718
 Qy 583 -----SerValGluLeuSerGlu ----- 588
 Db 2719 TCTGACCCGACCCGAGTCTTCTTCGGAGCGGTGTCGAGAACCCCGGCGCGCTGGGG 2778
 Qy 589 -----GlnSerLeuGlnGluLeuGlnAlaValLysArgAlaLysPro 602
 Db 2779 CTGGAGTCTGCTGGCTCGGAGCGCCCGGAGCCCTCGGGGCGACAGCGAGCGGAGACC 2838
 Qy 603 AspLeuValIleThr -----HisProAla 610
 Db 603 AspLeuValIleThr -----HisProAla 610

QY	166	ValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCysArgLeuAlaArgGlu	185
DB	187	---ACCAAATGGAACCCGGGTTCAGAGAGAGCTGCAGATGCTGTGTGCCCTTGGCCCGGAG	243
QY	186	GlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlnGlnLeuGluLeuArg	205
DB	244	GGCATCTCTGAAGCATCAAGCACAGCTTCTCAGAAAAGGACCTGAGAGATTGAAGCTTCAG	303
QY	206	GlySerLysValGlnThrLeuPheLeuSerLysGluLeuProGlyValLeuGluThr	225
DB	304	GGTTCCCAAGTTCAGACAATGTTCTCTCAGCAAGAACCCAGCTGCCAGAGTGCTAGAACT	363
QY	226	GluValThrTyrGlnPheIleAspGlnSerPheGlnPheLeuAlaLeuSerTyr	245
DB	364	GTGGTCACTACCACCAAGTTCATGTGACCAAGAGTTCAGAGATTCTTGCTGCTATGTATAC	423
QY	246	LeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGlyValGlyThrLeuLeuArg	265
DB	424	CTACTAGACGCTGAGGAGAGCCCCAGGAACTCCGAGGAAGTGTGCAGATGCTCTGCTC	483
QY	266	GlyAspAlaGlnProHisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeu	285
DB	484	TCTACGCGGGGCTGCGTGTGTCATCTGCCTACTCACCACTAGATTCTCTTTGACATGCTA	543
QY	286	SerAlaGluArgMetArgAspIleGluArgHisPheGlyCysMetValSerGluArgVal	305
DB	544	AGTACAGAGAGGATTCGTGACATTGGAAACCATTTGGCTGTGTGTGCCAGGGCTGTG	603
QY	306	LysGlnGluAlaLeuArgTTPValGlnGlyGlnGlyGlnGlyCysProGlyValAla	324
DB	604	AAACAGGACACCTTTCGCGTGGTACAAGGACAAAGCCAA-----CCCAAGGTGGGACA	657
QY	325	-----ProGluValThrGluGlyAlaLysGlyLeuGluAspThrGluGluGlu	342
DB	658	GTACGACGACGAAGAAGAGTAGCTCGAGGACGAGGAACAGAGGAGGAGAGAGAG	717
QY	343	GluGluGluGlyGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyrGluThr	362
DB	718	GAAGAAGAGGAGGAGGAACCACTTTGGACTGGAGCTGTGTACTGCCTGTATGAGACA	777
QY	363	GlnGluAspAlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArg	382
DB	778	CAAGAGGATGATTTGTTCGCCAGGCTCTCAGCAGCCTTCAGAGATGTTACTTGGACGGA	837
QY	383	ValArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCysProAla	402
DB	838	GTTAGGCTGACCCGATGGACCTTGAGTCTGAGCTACTCGGTGCAGTCTGCCCGGAC	897
QY	403	GlyGlnAlaLeuArgLeuIleSerCysArgLeuValAlaAlaGlnGlu-----LysLys	422
DB	898	GGCCAGGCTCTGAGACCTGTGTGAGCTGTGAGCTGTGCGGCAAGGAGAGAAGAAGAAG	957
QY	421	LysLysSerLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGlyThr	440
DB	958	ANGAAGACTTCATCAACCCGCTGAG-----GGTTCTCAAGACACC	999
QY	441	ThrLysGlnLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeu	460
DB	1000	GGGAACAACCCCGAGCTCTCTGCTGCTCCACTCTGTGAGGCAATGATTATCCAGCAA	1059
QY	461	CysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAsp	480
DB	1060	TGTGGTCTGAGTATTCTGACCTTGTACACTTGTGCAAACTCCCTGACAGTTTGTGGAGAC	1119
QY	481	LeuSerGluAlaLeuArgAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArg	500
DB	1120	CTTCTGAGGCTCTGAAGCTAGCTCTCTCCCTTAAGGAGGCTGGGCGCTCTCCAGAACCGA	1179
QY	501	LeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaThrProGlnCysArgVal	520
DB	1180	CTCACTGAGGAGGCGTCTGCTTTTACTAAGCAACAGCGCTGGCTTGGCCCAATGCAAGGTG	1239

Qy	521	GlnThrValArgValGlnLeuProAspProGlnArgGlyLeuGlnTyrLeuValGlyMet	540
Ds	1240	CAGACACTCAGGATACAGATGCCTGGGCTCCAAAGAGGTGATCCACTACCTGGTCAITGTG	12399
Qy	541	LeuArgGlnSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaPro	563
Ds	1300	CTCCAGCAGAGCCAGTCCTTAACCATCTTGACCTCAGTGGCTGTCACTGCTGGGACT	13599
Qy	561	MetValThrTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeuSer	590
Ds	1360	GTGGTGGAACTCTGTGTTCCAGGCCCTGAAACACCCCTAAATGTGSCCTAAAGACCCCTCAGT	14119
Qy	581	LeuAlaSerValGlnLeuSerGluGlnSerLeuGlnGlnLeuAlaValLysArgAla	600
Ds	1420	CTGACTCTCTGGAGCTGACTGAGATCCACTGAGAGAGCTTCAGCTGTGAAGACATTA	14799
Qy	601	LysProAspLeuValIleThrHisProAlaLeuAspGlyHisProGlnProProLysGlu	620
Ds	1480	AAGCCAGACTGGGCCATCATACATTCATAAATTTGGGCACACATCCTCAGCCTCTGAAGGGA	15399
Qy	621	LeuIleSerThrPhe	625
Ds	1540	TGAAGCAGTGTCTTC	1554
RESULT 10			
AAK83313			
IC	AAK83313 standard; DNA; 933 BP.		
AC	AAK83313;		
DT	07-NOV-2001 (first entry)		
XX	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38125.		
DE	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;		
KW	cytostatic; gene therapy; vaccine; metastasis; ds.		
OS	Homo sapiens.		
PN	WO200157182-A2.		
XX	03-AUG-2001.		
PF	17-JAN-2001; 2001WO-US01354.		
PR	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-C184664.		
PR	02-MAR-2000; 2000US-C186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		

PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226886.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241231.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244637.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 31-DEC-2000; 2000US-0250160.
PR 31-DEC-2000; 2000US-0250391.
PR 35-DEC-2000; 2000US-0251030.
PR 35-DEC-2000; 2000US-0251988.
PR 35-DEC-2000; 2000US-0256719.
PR 36-DEC-2000; 2000US-0251473.
PR 38-DEC-2000; 2000US-0251856.
PR 38-DEC-2000; 2000US-0251868.
PR 38-DEC-2000; 2000US-0251869.
PR 38-DEC-2000; 2000US-0251989.
PR 38-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 35-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SC: INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure; SEQ ID NO 38125; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells, AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 933 BP; 146 A; 309 C; 325 G; 153 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1.36e-98 Length: 933
Score: 1243.00 Matches: 245
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0


```
Query Match: 38.63% Indels: 0
DB: 22 Gaps: 0
US-10-029-347-2 (1-625) x AAK83313 (1-933)
Qy 1 MetLeuKlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGlnLeuProAla 20
Db 197 ATGTCGCCCGCAGCGCAGCGCTCTCTTCATCTTGACGCGCGCAGCGAGCTGCGCGG 256
Qy 21 LeuGlyGlyProGluAlaAaProCysThrAspPheGluAlaAlaSerGlyAlaArg 40
Db 257 CTGGGGCGCGCGCGCGCGCTGACAGACACCCCTTCAGGGCGGAGCGCGCGCGG 316
Qy 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuLeuValThrTr 60
Db 317 GTGCTAGCGGGCTGTGAACAAGGCGCTGCTGCCACGCGCCCTCTGCTGTGACCAAG 376
Qy 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
Db 377 CGCGCCCGCCCGCGGAGGCTGCGAGGCGCGCTGTGTTCCCGCAGGTGCGCGCAGG 436
Qy 81 ArgGlyPheSerAspLysAspLysLysLysLysLysPheTyrPheTyrPheArgAspGluArg 100
Db 437 CGCGCTTCTCCGACAGGACAGAGAGACTATTTCTACAGATTCTCCGGGATGAGAGG 496
Qy 101 ArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
Db 497 AGGGCGGAGCGCGCTACCGCTTCGTGAAGAGGAGACGAGACGCTGTTCGCGCTGT 556
Qy 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
Db 557 GTCCCTTCTGTGTGTGAATGTTGTGACCGCTGCTGCGCAGCAGCTGAGCTCGGTC 616
Qy 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrLeuLeuPheIleThrSerVal 160
Db 617 GACCTGTGCGGCGAGCTCAAAGACACACCGCTCAGTGTACCTGCTTTTCATCACCAG 676
Qy 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGluGlyAspLeuArgAsnLeuCys 180
Db 677 CTGAGCTCGGCTCCGGTAGCGAGCGGCGCGGCTTCAGGGCGAGCTGCGCATCTGT 736
Qy 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
Db 737 CGCGTGGCGCGGAGGCGCGCTCTCGAGCGAGCGCGCAGTTTCCCGAGAGGAGTGG 796
Qy 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
Db 797 CAACCTGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTTTCCAGCAAAAGGAGCTGC 856
Qy 221 GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPheLeu 240
Db 857 GCGGTGTGTGAGACAGAGGTCACCTACCACTTCATCGACGAGAGCTTCCAGAGAT 916
Qy 241 AlaAlaLeuSerTyr 245
Db 917 GCGGCACTGTCCAC 931
RESULT: 11
AAK6274:
ID AAK62741 standard; cDNA; 989 BP.
XX
AC AAK62741;
XX
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:7801.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
KW cytosolic; gene therapy; vaccine; metastasis; ss.
XX
OS Homo sapiens.
XX
PN W0200157182-A2.
```


D6 677 CTGAGCTCGCTCCGGTACCGACAGGCGCCCGCTTGCAGGGCGACCTGCGCAATCTGTGC 736
 QY 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
 D6 737 CGCTGGCCCGGAGGCGCTCTCGGACGCGCAATTTGCCGARAAGCACTGGAG 796
 QY 201 GlnLeuGlnLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuP 220
 D6 797 CAACCTGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTCTTCACCAAAAGGACCTGC 856
 QY 220 roGlyValLeuGluThrGluValThrTyGlnPheileAspGlnSerPheGlnGluP 239
 D6 857 CGGCGCTGTGGANACAGAGCTCAGTACAGTTATCGACCAAGAGCTTNCAGAGTT 916
 QY 239 eLeuAlaAla--LeuSerTyrrLeuLeu---GluAspGlyGlyValProArgThrAla 257
 D6 917 CTTCCGGGCGACCTGTCTCTACCTTGTGGAAGACGCGGCGGTGGCCCAAGAACCCGCC 976
 QY 258 Gly 258
 D6 977 GGN 979
 RESULT 12
 ABL90582
 ID ABL90582 standard; cDNA; 990 BP.
 XX ABL90582;
 AC ABL90582;
 XX
 DT 24-MAY-2002 (first entry)
 XX Human polynucleotide SEQ ID NO 1144.
 DE
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.
 XX Homo sapiens.
 OS
 XX W0200190304-A2.
 FN
 XX 29-NOV-2001.
 PC
 XX 18-MAY-2001; 2001W0-US16450.
 PF
 XX 19-MAY-2000; 2000US-205515P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PA Barse CE, Rosen CA;
 XX
 PI
 XX WPI; 2002-122018/16.
 DR
 XX P-PSDB; ABB90173.
 XX
 PT Note: 1435 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 XX C:aim 4; SEQ ID NO 1144; 2081pp + Sequence Listing; English.
 PS
 XX The invention relates to novel genes (AB09449-ABL90853) and proteins
 CC (AB09040-AB090444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 990 BP; 158 A; 318 C; 344 G; 163 T; 7 other;
 Alignment Scores:
 Pred. No.: 2 98e-90 Length: 990
 Score: 1147.50 Matches: 246
 Percent Similarity: 93.23% Conservative: 2
 Best Local Similarity: 92.48% Mismatches: 10
 Query Match: 35.66% Indels: 8
 DB: 24 Gaps: 1
 US-10-029-347-2 (1-625) x ABL90582 (1-990)
 QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheileLeuAspGlyAlaAspGluLeuProAla 20
 D6 197 ATGCTGGCCCGGAGCGCGCTCTTCATCTCGACGCGCGGACGAGCTGCCGGG 256
 QY 21 LeuGlyGlyProGluAlaAlaProCysThrAspPheGluAlaAlaSerGlyAlaArg 40
 D6 257 CTGGGGGGGCGCGAGCGCGCTTCACAGACCCCTTCAGGCGCGAGCGCGCGG 316
 QY 41 ValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThrThr 60
 D6 317 GTCTAGGCGGGCTGTGAGTAAGGCGCTGTGCCACCGCCCTCTCTGTGTGACACG 376
 QY 61 ArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGluVal 80
 D6 377 CGCGCGCGCGCGCGCGAGGTGCAGGGCGCGCTGTGTTCGCCGAGTGCAGCGAGGTG 436
 QY 81 ArgGlyPheSerAspLysAspLysLysLysTyrrPheTyrrLysPhePheArgaspGlyArg 100
 D6 437 CGCGGCTCTTCGACAGGACAAAGAAAGATATTTCTCAAGTWTTCGCGGAAGAGG 496
 QY 101 ArgAlaGluArgAlaTyrrArgPheValLysGluAsnGluThrLeuPheAlaLeuCysPhe 120
 D6 497 AGGCGCGAGCGCGCTACCGCTTCGTGAAGGACAGACGAGCTGTTCGCGCTGTCTTC 356
 QY 121 ValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGlyArg 140
 D6 557 GTGCCCTTCGT 616
 QY 141 AspLeuSerArgThrSerLysThrThrThrSerValTyrrLeuLeuPheileThrSerVal 160
 D6 617 GACCTGTGGCGCAGTCCAGACACACACGTCAGTGTACCTGTCTTTCATCACCGGTT 676
 QY 161 LeuSerSerAlaProValAlaAspGlyProArgLeuGlnGlyAspLeuArgAsnLeuCys 180
 D6 677 CTGAGCTCGGCTCCGCTAGCGCGCGCGCGCTTCAGCGCGGACCTGCACAACTGTGTGC 736
 QY 181 ArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 200
 D6 737 CGCTGGCGCGGAGGCGCTCTCGGACCGACGCGCAATTCGCCRAAAGAGACCTGGAG 796
 QY 201 GlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeu 219
 D6 797 CAACCTGAGCTTCGTGGCTCCAAAGTGCAGACGCTGTGTCTTCACCAAAAGGAGCTGC 856
 QY 220 ProGlyValLeuGluThrGluValThrTyrrGlnPheileAspGlnSerPheGlnGluP 239
 D6 857 CGCGGCGCTGTGGANACAGAGGTCACTACCACTTCATCGACCAAGAGCTTNCAGAGT 916
 QY 239 heLeuAlaAla--LeuSerTyrrLeuLeu---GluAspGlyGlyValProArgThrAla 257
 D6 917 TCTCTCGGCGGACCTGTCTTACCTTGTGTGAGGACGCGGCGGTGGCCCAAGAACCCGC 976
 QY 257 ag.y 258

```

Db          977 CGGN 980
RESULT 13
ID AA:47129 standard; DNA; 3300 BP.
XX
AC AA:47129;
XX
DT 20-AUG-2002 (first entry)
XX
DE Pyrin domain containing protein NALP3/PY5-bs coding sequence.
XX
KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW neuroprotective; antiaerobic; antirheumatic; antiaesthetic;
KW nephrotropic; osteoparatic; nocropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX
OS Unidentified.
XX
DN WC200240668-A2.
XX
PD 23-MAY-2002.
XX
PF 30-CCT-2001; 2001WO-EP12545.
XX
PR 15-NOV-2000; 2000DE-1056687.
PR 30-NOV-2000; 2000DE-1059595.
XX
FA (APOT-); APOTRCH RES & DEV LTD.
XX
PI Tschopp J, Martinon F;
XX
DR WPI; 2002-427093/45.
DR P-PSDB; AA017857.
XX
PT New DNA encoding protein with pyrin domain, useful for treating
PT diseases involving impaired signal transduction, particularly
PT inflammation, also proteins and antibodies
XX
PS Claim 5; Fig 1; 116pp; German.
XX
CC The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PYD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention.
XX
SQ Sequence 3300 BP; 726 A; 943 C; 979 G; 652 T; 0 other;

Alignment Scores:
Score: 1.32e-61 Length: 3300
Score: 826.50 Matches: 23
Percent Similarity: 48.00% Conservative: 105
Best Local Similarity: 33.00% Mismatches: 247
Query Match: 25.68% Indels: 119
DB: 24 Gaps: 17

US-10-029-347-2 (1-625) x AAL47129 (1-3300)
QY 1 MetLeuAlaGlnProGlnArgLeuPheLeuAspGlyAlaAspGluLeu---Pro 19
Db 763 CTCATCCGAGTTCGCGAGCGCTCTCTTTCATCATCGAGCGCTTCGATGAGCTCAAGCCT 822
QY 20 AlaLeuGlyGYPProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAla 39

```

```

Db          823 TCTTTCCAGATCCTCAGGACCCCTGCTGCTCTGCGAGGAGAAACGGCCACGGAG 882
QY 40 ArgValLeuGlyGlyLeuLeuSerIysAlaLeuLeuProThrAlaLeuLeuValThr 59
Db 883 CTGCTTCTTAACAGCTTAATTTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCAC 942
QY 60 ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu 79
Db 943 ACACGCCCCACGGCTTTGGAGAAGCTCCACCGTCTGCTGGAGCACCCAGGATGTGGAG 1002
QY 80 ValArgGlyPheSerAspIysIysIysIysIysIysIysIysIysIysIysIysIysIys 99
Db 1003 ATCTCTGGCTTCTCTCAGGACAGAAAGGAAGTAATCTCTACAACTATTTCACATGCA 1062
QY 100 ArgArgAlaGluArgAlaArgPheValIysGluAsnGluThrLeuPheAlaLeuCys 119
Db 1063 GAGCAGCGGGCCAGTCTTCAATTACGTGAGGACAAAGACCTCTCTTCCACATGTGC 122
QY 120 PheValProPheValCysTrpIleValCysThrValLeuArgGlnGlnLeuGluLeuGly 139
Db 123 TCTGCTCCCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1482
QY 140 ArgAspLeuSerArgThrSerIysThrThrSerValIysIysIysIysIysIysIys 159
Db 1183 GGGCTGTTGAGACAGACGCTCCAGGACCACTGTCAGTGTATCTCTTACCTGCTGAGT 1242
QY 160 ValLeuSerSerAlaProValAlaAspGlyProArgLeuGln-----GlyAspLeuArg 177
Db 1243 CTGATGCAACCAAGCGGGGGCC-----CGCGCTTCCAGCCCCCACCACACAGAGA 1296
QY 178 AsnLeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluIys 197
Db 1297 GGGTTGTGCTCTCTGCGGCAGATGGCTCTGGAATCAGAAATCTATTTTTCAGGAGCAG 1356
QY 198 GluLeuGluGlnLeuGluLeuArgGlySerIysValGlnThrLeuPheLeuSerIysIys 217
Db 1357 GACCTCCGGAAGCAGCGCTAGACGGGGAAGACGTC---TCTGCTTCTCTCAACATGAAC 1423
QY 218 GluLeuProGlyValLeuGluThrGluValThrIleGlnPheIleAspGlnSerPheGln 237
Db 1414 ATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCTCATCTTGAGTTTCCAG 1473
QY 238 GluPheLeuAlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAla 257
Db 1474 GAATCTTTGACGATGTACTATATCTCGACGAGGGG-----GAG 1515
QY 258 GlyGlyValGly-----ThrLeuLeuArgGlyAspAlaGlnPro 270
Db 1516 GCGGGGCGAGGCCACAGACCGAGCTGACCGGCTGTGTACCGAGTACGCGTCTTCTGAA 1575
QY 271 HisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMet 290
Db 1576 AGGAGCTTCTGGCACTCACAGCGCTTCTGTTTGGACTCTCTGACGAGGAGCAGG 1635
QY 291 ArgAspIleGluArgHisPheGlyCysMetValSerGluArgValIysGlnGluAlaLeu 310
Db 1636 AGCCACCTCGAGAGAGTCTCTGCTGGAAGGTCTCTCGCGCACATCAAGATGACCTGTG 1695
QY 311 ArgTrpValGlnGlyGlnGlyCysProGlyValAlaProGluValThrGluGly 330
Db 1696 CAGTGGATCCAAAGCAAGCTCAG----- 1719
QY 331 AlaIysGlyLeuGluAspThrGluGluProGluGluGluGluGluGluGluGluProAsn 350
Db 1720 -----AGCCAGCGCTCCACCTCGACGAGGGCTCC----- 1749
QY 351 TyrProLeuGluLeuLeuTyrCysLeuTyrGluThrGlnGluAspAlaPheValArgGln 370
Db 1750 -----TTGGAGTCTTTCAGCTGCTGTACGAGATCCAGGAGGAGGAGTTTATCCAGCAG 1803
QY 371 AlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPheCysArgMetAspVal 390
Db 1804 GGCCTGAGCCACTTCCAGTGTATGCTGCTCAGCAACATTGCC---TCCAAGATGGAGCAC 1860

```

```
QY 391 AlaValLeuSerTyrCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuIleSer 410
DB 1861 ATGGCTCTCTCTTCTGTTGAGGCTGCAGAGCGCCAGTCTGCTGCACTTGTATGCG 1920
QY 411 CysArg----- 412
DB 1921 GCCACCTACAGCGCGAGCGGAGACCGCGGAGGTCTCCGAGGAGCGCACACGCTG 1980
QY 413 LeuValAlaAlaGlnGluLysLysSerLeuGlyLysArgLeuGlnAlaSerLeu 432
DB 1981 TTGGT-GCAGCTCAGACACAGAGGACCGTCTGCTGGA---CGCTACAGTGAACATCT 2036
QY 433 GlyGlyGly-----SerSerGlnGlyThrThrLysGlnLeuProAla-SerLe 448
DB 2037 GGCAGCGGCGCTGTGCACCACTCCAAACCTGTATAGAGCTGTCTGTGTACCGAAATGCCCT 2096
QY 448 u-----LeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLeuSe 464
DB 2097 GGGCAGCGCGGGGTGAAGCTGCTCTGTCAGGACTCAGACACCCCACTGCAACTTCA 2156
QY 464 rSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerGluAl 484
DB 2157 GAACCTGAGGCTGAAGAGGTGGCGATCTCCAGCTCAGCTCGAGGACCTCTCTGCASC 2216
QY 484 AlaArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGluAl 504
DB 2217 TCTCATAGCAATAGAAATTGACAGAGTGGATCTCAGTGCACAGCGGTGGATCTCC 2276
QY 504 aGlyLeuArgMetLeuSerGlnGlyLeuAlaTrpProGlnCysArgValGlnThrValAr 524
DB 2277 AGGCATGATGCTCTTTGCGAGGGCTGGCGCATCCCACTGAGCTGAGGATGATTC 2336
QY 524 gValGlnLeuProAspProGlnArgGly--LeuGlnTyrLeuValGlyMetLeuArgGl 543
DB 2337 GTTGAGGAAGTGTGACAGCTGGAGTCCGGGCTTGTTCAGGAGATGGCTTCTGTGCTG 2396
QY 543 nSerProAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValTh 563
DB 2397 CAACCCACATCTGTTGAGTGGACCTGACAGAAATGCACTGAGGATTTGGGCTGAG 2456
QY 563 rTyrLeuCysAlaValLeuGlnHisGlnGlyCysGlyLeuGlnThrLeu----- 579
DB 2457 GTTACTATGCGAGGACTGAGGACCCAGCTCTGCAGACTACGAGCTTTGTGCTGAGAT 2516
QY 580 -----SerLeuAlaSerValGluLeuSerGluGlnSe 590
DB 2517 CTGCGGCTCAGTCTGCTGCTGTGACGAGCTGGCTCACTCTCAGTGTGAACCGAG 2576
QY 590 rLeuGlnGluLeuGlnAlaValLysArgAlaLysProAspLeuValIle----- 606
DB 2577 CTGAGAGAGCTGGACCTGAGCTGATGAGCTGGGGGACCTCGGGGTGCTCTGCTGTG 2636
QY 607 -----ThrHisPro----- 609
DB 2637 TGAGGSCCTCAGGCACTCCACGTCGCAAGCTCCAGACCTCGCGTGGCATCTCGGCT 2696
QY 610 -----AlaLeuAspGlyHisProGlnProPro 618
DB 2697 GGGCTCTGCGGCTGTGAGTGTCTTTCTGTGTGTCTCA-GGCCAACCAACACCTCCG 2753
RESULT 14
ID AAL44363
AC AAL44363;
XX
XX
XX 31-OCT-2002 (first entry)
XX Human PYRIN-8 cDNA sequence #2.
XX Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory response; inflammatory disorder;
```

```
KW immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW leukemia; autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-8.
OS Homo sapiens.
PH Key Location/Qualifiers
FT CDS 1..3186
FT /tag= a
FT /product= "Human PYRIN-8 protein #2"
XX
XX W0200261049-A2.
XX 08-AUG-2002.
XX 31-JAN-2002; 2002WO-US02967.
XX 31-JAN-2001; 2001US-265231P.
XX 20-SEP-2001; 2001US-318645P.
XX (MILL-) MILLENNIUM PHARM INC.
XX (AMHP ) WYETH.
XX Bertin J, Wang W, Blatcher M;
XX WPI: 2002-627477/67.
XX P-PSDB; AAO15590.
XX New PYRIN polypeptides and nucleic acids useful for modulating and
XX diagnosing stress-related, apoptotic and inflammatory responses, or for
XX treating inflammatory and immune system disorders, cancers, or
XX neurological diseases -
XX Claim 4; Fig 8; 167pp; English.
XX The invention comprises the amino acid and coding sequences of human
XX PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
XX useful for modulating and diagnosing stress-related, apoptotic and
XX inflammatory responses. The PYRIN protein and DNA sequences are useful
XX for treating inflammatory disorders and immune system disorders (e.g.
XX Crohn's disease, reactive arthritis, multiple sclerosis, contact
XX dermatitis, psoriasis, graft rejection, allergies, viral infections and
XX bacterial infections); cancer (e.g. leukemia); autoimmune disorders
XX (e.g. systemic lupus erythematosus and arthritis); and neurological
XX diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
XX protein and DNA sequences may also be used in screening assays, detection
XX assays (e.g. chromosomal mapping, tissue typing or forensic biology),
XX predictive medicine (e.g. diagnostic assays, clinical trials and
XX pharmacogenomics) and transcription profiling. The present DNA sequence
XX encodes a human PYRIN-8 protein.
SQ Sequence 3186 BP; 701 A; 908 C; 945 G; 632 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 1..54e-61 Length: 3186
Score: 825.50 Matches: 230
Percent Similarity: 48.07% Conservative: 106
Best Local Similarity: 32.90% Mismatches: 247
Query Match: 25.65% Indels: 118
DB: 24 Gaps: 17
```

US-10-029-347-2 (1-625) x AAL44363 (1-3186)

```
QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeu---Pro 19
DB 844 CTCATCCGAGTTCGAGCGCTCTTTTCATCATCGCGCTTCGATGAGCTCAAGCCT 903
QY 20 AlaLeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSerGlyAla 39
DB 904 TCTTCCACGATCTCAGGAGCCCTGGTGCCTCTGTGGAGGAGAAACGGCCCGAG 963
```


KW glomerulonephritis; neurological disorder; Alzheimer's disease; ALS;
 KW amyotrophic lateral sclerosis; Parkinson's disease; Huntington's disease;
 KW aplastic anaemia; ischaemia; meningitis; liver disease; Crohn's disease;
 KW insulin-dependent diabetes; multiple sclerosis; Grave's disease; HIV;
 KW human immunodeficiency virus; tuberculosis; lepromatous leprosy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH 139..3243
 FT CDS /*tag= a
 FT /note= "Human PYRIN-1 protein"

XX W0200161805-A2.

XX 23-AUG-2001.

XX 16-FEB-2001; 2001WO-US40145.

XX 17-FEB-2000; 2000US-0506087.

PR 01-SEP-2000; 2000US-0653901.

XX (MILL-) MILLENNIUM PHARM INC.

XX Bertin J;

XX WPI; 2001-514773/56.

DR P-PSDB; AAEC7514.

XX Novel polypeptide comprising NBS-1 or PYRIN-1 of the pyrin domain
 PT protein family useful in screening and detection assays and for
 PT treating, e.g., cancer, viral infections, autoimmune disease, and
 PT Alzheimer's.

XX Claim 1; Fig 4; 11pp; English.

XX The invention relates to human NBS-1 nucleotide binding site) and
 CC PYRIN-1 protein and nucleic acids. NBS-1 and PYRIN-1 DNA, protein and
 CC their modulators are useful in the treatment of apoptotic and
 CC inflammatory disorders, cancer (leukaemia, melanoma, carcinoma); viral
 CC infections (including herpesvirus and adenovirus), autoimmune diseases
 CC (systemic lupus erythematosus (SLE), immune-mediated glomerulonephritis,
 CC arthritis; neurological disorders (Alzheimer's disease, Parkinson's
 CC disease, amyotrophic lateral sclerosis (ALS), retinitis pigmentosa,
 CC Huntington's disease); aplastic anaemia, ischaemia, meningitis, liver
 CC diseases. NBS-1 and PYRIN-1 DNA, protein and their modulators are also
 CC used for the treatment of inflammatory and immune disorders such as
 CC chronic inflammatory diseases such as Crohn's disease, insulin-dependent
 CC diabetes, organ specific autoimmunity, including multiple sclerosis,
 CC Grave's disease, allergy, asthma, HIV, tuberculosis and lepromatous
 CC leprosy. The present sequence is a cDNA encoding human PYRIN-1 protein.

XX Sequence 3857 BP; 958 A; 931 C; 1023 G; 945 T; 0 other;

SQ Alignment Scores:

Pred. No.:	3e-59	Length:	3857
Score:	800.50	Matches:	219
Percent Similarity:	43.64%	Conservative:	117
Best Local Similarity:	28.44%	Mismatches:	241
Query Match:	24.88%	Indels:	193
DB:	22	Gaps:	17

US-10-029-347-2 (1-625) x AAD:4323 (1-3857)

QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20

Db 1000 ATCGTGAGAAACCCCTCCAGATCCTCTCCATGAGCGGCTTCGATGACCTCAAGGT 1059

QY 2: LeuGlyGlyProGluAlaAlaPro---CysThrAspProPheGluAlaAlaSer-GlyAla 39

Db 1060 GCCTTTGACGACACATAGGACCGCTCTGCATCTGACGTGGCAGAGCGCGGAGAC 1119

QY 40 ArgValLeuGlyGlyLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 59

Db	1120	ATTCCTCTGACGACCTCATCAGAAAGAGCTGCTCCCGAGGCTCTCTCTGCTGACACC	1179
QY	60	ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAlaGlu	79
Db	1180	ACGAGACCTGTGGCCCTGGAGAAATCGCAGCATTGCTGGACCATCTCGGCATGTGAG	1239
QY	80	ValArgGlyPheSerAspCysAspLysLysLysLysPheThrLysPheArgAspGlu	99
Db	1240	ATCTCTGGCTTCTCCGAGGCCAAAGAAAGAGTACTTCTTCAAGTACTTCTTCTGATGAG	1299
QY	100	ArgAAGAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAlaLeuCys	119
Db	1300	GCCCAAGCCAGCCAGCCTTCAGTGTGATTGAGGAGACGAGGCTCTTCCACATGTGC	1359
QY	120	PheValProPheValCysTrpIleValCysThrValLeuAAGlnGlnLeuGluLeuGly	139
Db	1360	TTCATCCCTCTGCTGCTGATCGTGTGCACTGGACTGAAACAGACAGATGAGAGTGGC	1419
QY	140	ArgAspLeuSerArgThrSerLysThrThrSerValTyrLeuLeuPheIleThrSer	159
Db	1420	AAGAGCCTTGCCAGACATCTAAGACCACCGCGGTGATGCTTCTTCTTCTTCCAGT	1479
QY	160	ValLeuSerSerAlaProValAlaAlaAspGlyProArgLeuGlnGlyAspLeuArgAsn	179
Db	1480	TTGCTGCAGCCCGGAGGAGCCAGGAGCCGCTCTCGCCGCCCTCTGGGGGCTC	1539
QY	180	CysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeu	199
Db	1540	TGCTCTTTGGCTGCAGATGGAATCTGGAACCAACCAAAATCTGTTTGGAGGATCCGACCTC	1599
QY	200	GluGlnLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeu	219
Db	1600	AGGAATCATGTGACTGCAGAGCGGATGTG---TCTGCTTCTGAGTGAACCTGTTC	1656
QY	220	ProGlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPhe	239
Db	1657	CAAAAGGAGTGGACTGCGAGAGTTCTACAGCTTATCCACATGACTTCTCCAGAGTTC	1716
QY	240	LeuAlaAlaLeuSerTyrLeuLeuGluAspGlyGlyValProArgThrAlaAlaGlyGly	259
Db	1717	TTTGGCCCATGTACTACTCTGCTGGAAGAAAAGGAAAGGAGGACCAACCTTCCAGG	1776
QY	260	-----ValGlyThrLeuLeuArgGlyAspAlaGln---	269
Db	1777	AGTCGTTTGAAGCTTCCAGCCGAGAGTGACAGCTTCTTGGAAACTATGCGAAATTC	1836
QY	270	ProHisSerHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArg	289
Db	1837	GAAGAGGGGTATTGATTTTCTGCTTTCTCTTTGGCTGCTGTAACACGAGGAGG	1896
QY	290	MetArgAspIleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAla	309
Db	1897	ACCTCTCTACTTGGAGAAAGAAATTAAGTTTGCAAGATCTCTCAACAACTCAGGTGGAGTG	1956
QY	310	LeuArgTrpValGlnGlnGlnGlyGlnGlyCysProGlyValAlaProGluValThrGlu	329
Db	1957	CTGAATGGATT-----GAATGAAAGCC	1980
QY	330	GlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGluGluGluGluGluPro	349
Db	1981	AAAGCTAAAGCTGCAG-----ATCCAGCCCGCCAG-----	2013
QY	350	AsnTyrProLeuGluLeuLeuTyrCysLeuTyrGluThrGlnGluAspAlaPheValArg	369
Db	2014	-----CTGGAATTTCTACTGTTTGTACGAGATGCGAGGAGGACTTCGTGCAA	2064
QY	370	GlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgValArgPheCysArgMetArg	389
Db	2065	AGGGCCATCGACTATTTCCTCAAGATTGAGATCAAT-----CTTCCACCAAGATGGAC	2118
QY	390	ValAlaValLeuSerTyrCysValArgCysCysProAlaGlyGlnAlaLeuArgLeuIle	409

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n mode:

Run on: October 3, 2003, 13:50:45 ; Search time 101 Seconds
(without alignments)
2731.331 Million cell updates/sec

Title: US-10-029-347-2

Perfect score: 3218

Sequence: : MAAQPORLLFIIDGADLPA.....ITHPALDGHQPQPKELISTF 625

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delep 6.0 , Delep 7.0

Searched: 569978 seqs, 22069:566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spo01/US10029347:/unat_02102003_173437_17096/app query.fasta_1.775
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFC1=0
-LOOPEXT=0 -UNIT5=bits -START1=1 -END1=1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029347 @CGN 1.1 56 @unat_02102003_173437_17096 -NCPJ=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:
2: /cgn2_6/prodata/2/ina/5A.COMB.seq:
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:
4: /cgn2_6/prodata/2/ina/65.COMB.seq:
5: /cgn2_6/prodata/2/ina/65.COMB.seq:
6: /cgn2_6/prodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	680.5	21.1	678	4	US-09-016-434-1011
2	400	12.4	257	4	US-09-016-434-208
3	325	10.1	4543	2	US-08-519-547A-5
4	312.5	9.7	4441	4	US-09-641-999-2
5	304.5	9.5	3333	1	US-08-295-502-1
6	304.5	9.5	3333	5	PCT-US95-10691-1
7	289	9.0	2859	4	US-09-099-041A-9
8	289	9.0	2859	4	US-09-245-281-9
9	289	9.0	2859	4	US-09-359B-9
10	289	9.0	2859	4	US-09-340-620A-9
11	289	9.0	3382	4	US-09-099-041A-7
12	289	9.0	3382	4	US-09-245-281-7

13	289	9.0	3382	4	US-09-207-359B-7
14	289	9.0	3382	4	US-09-340-620A-7
15	285	8.9	4302	4	US-09-245-281-38
16	285	8.9	4302	4	US-09-207-359B-38
17	285	8.9	4302	4	US-09-340-620A-38
18	251.5	7.8	4141	4	US-09-245-281-42
19	251.5	7.8	4141	4	US-09-207-359B-42
20	251.5	7.8	4141	4	US-09-340-620A-42
21	220.5	6.9	32042	4	US-09-245-281-44
22	220.5	6.9	32042	4	US-09-340-620A-63
23	216	6.7	1371	2	US-08-910-731-1
24	216	6.7	1371	2	US-08-795-395-1
25	214	6.7	1371	2	US-08-910-731-7
26	212	6.6	1386	2	US-08-910-731-5
27	209.5	6.5	1374	2	US-08-910-731-3
28	209.5	6.5	1374	2	US-08-795-395-3
29	154.5	4.8	3080	4	US-09-099-041A-25
30	154.5	4.8	3080	4	US-09-245-281-25
31	154.5	4.8	3080	4	US-09-207-359B-25
32	154.5	4.8	3080	4	US-09-340-620A-25
33	151	4.7	1470	4	US-09-099-041A-27
34	151	4.7	1470	4	US-09-245-281-27
35	151	4.7	1470	4	US-09-207-359B-27
36	137.5	4.3	3116	4	US-09-340-620A-27
37	137.5	4.3	3116	4	US-09-904-615-43
38	132.5	4.1	1876	4	US-09-620-405B-477
39	132.5	4.1	1876	4	US-09-604-287A-477
40	123	3.8	2072	4	US-09-016-434-1181
41	120	3.7	2346	3	US-09-193-503B-5
42	120	3.7	2346	4	US-09-415-839-5
43	117.5	3.7	2196	4	US-09-252-991A-2198
44	117.5	3.7	2196	4	US-09-620-405B-484
45	117.5	3.7	1518	4	US-09-604-287A-484

ALIGNMENTS

RESULT 1

US-09-016-434-1011
; Sequence 1011, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555

```
TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 678 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINCT04
; CLONE: 927003
US-09-016-434-1011

Alignment Scores:
Pred. No.: 5,19e-62 Length: 678
Score: 680.50 Matches: 77
Percent Similarity: 81.28% Conservative: 1
Best Local Similarity: 80.82% Mismatches: 36
Query Match: 21.5% Indels: 1
DB: 4 Gaps: 0

US-10-029-347-2 (1-625) x US-09-016-434-1011 (1-678)
Cy 108 PheValLys-GluAsnGluThrLeuPheAlaLeuCysPheValProPheValCysTrpI 127
Db 36 TCGTGAAGCGCTCTTANACCATANCNCTGTCTTCGTGCCCTTCGTGCTGGAT 95
Cy 127 eValCysThrValLeuArgGlnGlnLeuGluLeuGlyArgAspLeuSerArgThrSerLy 147
Db 96 CGTGTGCACCGTCTGCGCCAGCACTCGGCTCGGACCTGTGCGGCAGTGTGCGCACTCAA 155
Cy 147 s-ThrThrThrSerValThrLeuLeuPheLeuThrSerValLeuSerSerAlaProVala 167
Db 156 GGACACACACGTCAGTGATCTCTTTTCATCACCANNNTTCTGAGCTCGGCTCGGTAG 215
Cy 167 laAspGlyProArgLeuGlnGlyAspLeuAAsnLeuCysArgLeuAlaArgGluGlyV 187
Db 216 CCGACGGCCCGGTTGCGAGCGCACCTGCGCAATCTGTGCGCTGCGCCGCGAGGCG 275
Cy 197 alLeuGlyArgArgAlaGlnPheAlaGluGlyGluLeuGluGlnLeuGluLeuArgGlyS 207
Db 276 TCCCGGACGCGAGCGCGGANTTTCGCGAGAGGAACTGGAGCAACTGGAGCTTCGTGGCT 335
Cy 207 erCysValGlnThrLeuPheLeuSerLysGluLeuProGlyValLeuGluThrGlyV 227
Db 336 CCAAGTGCAGACGCTGTCTTCANCAABAAGGA-CTCGCGGCGGTGCTGGAACAGAGG 394
Cy 227 alThrTyGlnPheLeu-AspGlnSerPheGlnGluPheLeu-AlaAlaLeuSerTyRde 246
Db 395 TCACCTACCACTTCATCGSACACAGAGCTTCNAGGAATTCCTCCGCGGCACCTGTCTACCT 454
Cy 246 uLeuGluAspGlyGlyValFroArgThrAlaAlaGlyGlyVal-GlyThrLeuLeuArgG 266
Db 455 GCTGGAAACGNGGGGTGCGCACTGACCGCGCTGCGCGGCTTTGGGACATCTCTGCTT 514
Cy 266 lYAspAlaGlnProHisSerHisLeuValLeuThrArgPheLeuPheGly-LeuLeuS 286
Db 515 GAGACGCCCA-CCGACACGCACTTGGTGC-CAC-ACCGCTTCTCTTCGACCTGCTGA 572
Cy 286 eAlaGluArgMetArgAspIleGluArgHisPheGlyCysMetValSerGluArgVal 306
Db 573 ANGTGAGCGAAG-CGNACATCNACCCCACTC-CGCTGATGATGTTCGAACNTTTTA 630
Cy 306 yaGln-GluAlaLeuArgTrpValGlnGlyGlnGlyCysPro 321
Db 631 ACCA-GAAGCCCTGCGGTGGGTTTCAGAACAGGANTGNTNCCCG 676

RESULT 2
US-09-016-434-208
; Sequence 208, Application US/09/16434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhammer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 208:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SINTBST01
; CLONE: 1499408
US-09-016-434-208

Alignment Scores:
Pred. No.: 2,71e-33 Length: 257
Score: 400.00 Matches: 81
Percent Similarity: 96.47% Conservative: 1
Best Local Similarity: 95.23% Mismatches: 3
Query Match: 12.43% Indels: 1
DB: 4 Gaps: 0

US-10-029-347-2 (1-625) x US-09-016-434-208 (1-257)
Cy 91 TyrPheTyLysPhePheArgAspGluArgAlaGluArgAlaTyArgPheValLys 110
Db 2 TATTTTCAAGTATTTTCGGGATGAGAGGAGCGCGCGCTTACCGCTTCGTGAAG 61
Cy 111 GluAsnGluThrLeuPheAlaLeuCysPheValProPheValCysTrpIleValCysThr 130
Db 62 GACAACGAGACGCTGTTCGGCTGTCTTCGTGCCCTTCGTGTGTGATCGTGTSCAC 121
Cy 131 ValLeuArgGlnGlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThr 150
Db 122 GTGCTCGGCCA-NANTNGAGCTCGGTGCGGACCTTCGCGCACGTCCAGACCCACG 180
Cy 151 SerValTyLLeuLeuPheIleThrSerValLeuSerSerAlaProValAlaAspGlyPro 170
Db 181 TCAGTGTACTGCTCTTTCACACACGCTTCGAGCTCGGCTCGGTCGCGACCGGCC 240
Cy 171 ArgLeuGlnGlyAsp 175
Db 241 CGGTTCAGGGCGAC 255

RESULT 3
US-08-519-547A-5
```

Sequence 5, Application: US/08519547A
 Patent No. 5994082
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: Proteins Essential for the Expression of
 TITLE OF INVENTION: Vertebrate MHC Class II Genes, DNA Sequences Encoding Same
 TITLE OF INVENTION: and Pharmaceutical Compositions
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FISH & NEAVE
 STREET: 251 AVENUE OF THE AMERICAS
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: U.S.A.
 ZIP: 10020-1104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WordPerfect 6.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/519,547A
 FILING DATE: 25-AUG-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: BP94115378.7
 FILING DATE: 26-AUG-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: HALEY, JAMES F.
 REGISTRATION NUMBER: 27,794
 REFERENCE/DOCKET NUMBER: VOS-11
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-596-9000
 TELEFAX: 212-596-9090
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4543 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEetical: NO
 ANTI-SENSE: NO
 US-08-519-547A-5

Alignment Scores:
 Pred. No.: 2,41e-23 Length: 4543
 Score: 325.00 Matches: 195
 Percent Similarity: 36.93% Conservative: 52
 Best Local Similarity: 26.00% Mismatches: 296
 Query Match: 10.10% Indels: 179
 DB: 2 Gaps: 29

US-10-029-347-2 (1-625) x US-08-519-547A-5 (1-4543)

QY	1	MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20	1808	GAGCTGTCCGGCTTCTCCATGAGCAGCGCCAGGCATACGTGATCGCTACTTTTGAGAGC 1867
Db	1574	ATCTTGAAGAGACCTGACCGGTTCGTCTATCTAGAGCCTTCGAGGAGCTGGAGCG 1633	99	GluArgArgAlaGlu-----ArgAlaTyArgPheValLysGluAsnGluThrLeu 115
QY	21	LeuGlyGlyProGluAlaProCys---ThrAspPheGluAlaAlaSerGlyAla 39	1868	TCAGGATGACAGACACCAAGACAGACCTCTCCGGGACCGGCCACTTCTT 1927
Db	1634	CAAGATGGCTTCTGACACAGCTGCGGACCGGCGGCGGCTCTCCCTCCCG 1693	116	PheAlaLeuCysPheValProPheValCysTrpIleValCysThrValLeuArgGln 135
QY	40	ArgValLeuGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 59	1928	CTCAGTCACAGCCACAGCCCTACTTTGTCCGGCAGTGTCAGCTCTCAGAGCCCTG 1987
Db	1694	GGCGCTGCGCGGCTTTTCAGAGAAGAGCTCTCCGAGGTTCACCTCTCTCTCA 1753	136	LeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyLeu 154
QY	60	ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla 78	1988	CTGGAGCTTGGGAGGAC---GCCAAGTCGCTCCACGCTCAGGGACTCTATGTCGC 2044
Db	1754	GCCCGCCCCGG-----GGCCGCTGTGTCCAGAGCCTGAGCAGCGCCAGCCCTATT 1807	155	LeuPheIleThrSerValLeuSerSerAlaProValAlaAspGlyProArgLeuGln 174
QY	79	GluValArgGlyPheSerAspLysLysLysTyPheTyLysPheArgAsp 98	2045	CTGTGGGCGCTGCAGCCCTCGACAGCCCCC-----GGG 2080
			175	AspLeuArgAsnLeuCysArgLeuAlaArgGluGlyValLeuGlyArgAlaGlnPhe 194
			2081	GCCCTGGCAGAGCTGGCCCAAGCTGGCTGGGAG-----CTGGGCGCAGACATCAA 2131
			195	AlaGluLysGluLeuGluGlnLeuArgGlySerLysValGlnThrLeuPheLeu 214
			2132	-----AGTACCCTACAGGAGGACAGTTCCCATCCGACAGCTGAGACCTGGGCGATG 2185
			215	SerLys-----LysGluLeuProGlyValLeuGluThrGluValThrTyr---Gln 230
			2186	GCCAAAGGCTTAGTCCAAACACCCCGGCGGCGCAGAGTCCGAGCTGGCTTCCCCAGC 2245
			231	PheIleAspGlnSerPheGlnPheLeu---AlaAlaLeuSerTyLeuLeuGluAsp 249
			2246	TTCTCTCTGCAATGCTTCTCGGGGCGCTGTGGTGGCTCTGAGTGGCGAAATCAAGAC 2305
			250	GlyGlyValProArgThrAlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGln 269
			2336	AAGGAGCTCCCGCATCTACCTAGCA-----TTGACCCCAAGGAAGAAGAGG 2550
			270	ProHisSerHis---LeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGlu 288
			2351	CCCTATGACAACTGGCTGGAGGGGTGCCACGCTTCTGGCTGGGCTGATCTTCAGGCT 2410
			289	ArgMetArgAspIleGluArgHisPheGlyCysMetValSerGluArgValLysGln 308
			2411	CCCGCCGCTGCTGGGAGCCCTACTCGGGCCATCGGGCGCTCGCTCGTGA-CAGAA 2463
			309	AlaLeuArgTrpValGlnGlyGlnGlyGlnGlyCys---ProGlyValAlaProGluVal 328
			2470	GCAGA-AGTGCTTGGAGGACCTGAAGCGGCTGCAGCCGGG-----A 2513
			328	hrGluGlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGluGlyGluG 348
			2514	CATCGCGGCGCGGAGCTG----- 2533
			348	luProAsnTyProLeuGluLeuLeuTyCysLeuTyGluThrGlnGluAspAlaPheV 368
			2534	-----CTTGAGCTGTGCTGACTCGGCCAGCGCCAGGAGCGGAGGCTGGAATT 2579
			368	alArgGlnAlaLeuCysArgPhePro---GluLeuAlaLeuGlnArgValArgPheCysA 387
			2580	GGCAGACGTGTGACAGAGCTCCCGGCGGCTCTCTTTCTGGGACCCCGCTCAGGC 2639
			387	rgMetAspValAlaValLeuSerTyCysValArgCys----- 399
			2640	CTCCTGATGCATGTACTGGGCAAGGCTTGGAGGGCGGCGGCAAGACTTCTCCCTGG 2699
			400	-----CysProAlaGly---GlnAlaLeuArgLeuLeuSerCys- 411
			2700	ACCTCCGACGACTGGCATTTGCCCTCTGTGATTTGGGGAGGCTCTGTGGGACTCAGCTGTG 2759
			411	----- 411
			2760	TCACCCGTTTCAGGGCTGCCTTGAGCGACAGCGTGGCGCTGTGGGAGTCCCTCGCGGAGC 2819

```
QY 412 -----ArgLeuValAlaAlaGlnGlyLys-----LysL 421
Db 2820 ATGGGGAGACCAAGCTACTTACGACAGAGAGAGATTCACGCTCGAGCCCTTCAAG 2879
QY 421 ySLysSer-----LeuGlyLysA:gluGln:AlaSerLeuGlyG 435
Db 2880 CCAAGTCCCTGAGGAGTGTGGAAGACCTGGGAAGCTTGTGCAACTCAGAGGACGAGAA 2939
QY 435 ySerSerGlnGlyThrThrLysGlnLeuProAla-----S 447
Db 2940 GTTCTCGGAGACACAGACTGGGAGCTCCCTGCTGTTCGGAGCTTAAGAAACTGGAGT 2999
QY 447 erLeuLeuHisProLeuPheAlaMetThrAspPro----- 459
Db 3000 TTGGCTGGGCCCCCTGTCTCAAGCCCGGAGCTTTCCTCCAAACTGGTGGGATCTCCACG 3059
QY 460 -----LeuCyHisLeuSerSerLeuThrLeuSerHisCySLysLeuProAspA 476
Db 3060 CTTTCTCCTCCCTGCAGCATTTGACCTGGATGGCTGCTGATGAGAACAGATCGGGAGC 3119
QY 476 laValCysArgAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyC 496
Db 3120 AGGCTGTCTCGCAGCTCTCAAGCCCTTCCCTCCAGCTGAAGTCTTGGAAACCTCAATC 3179
QY 496 euLeuHisAspArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpP 516
Db 3180 TGTCCAGAACACATCACTACCTGGGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3233
QY 516 roGlnCysArgVa:GlnThrValArgValGln----- 526
Db 3234 CTTCGCTCGTGCATCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3293
QY 527 -----LeuProAsp----- 529
Db 3294 GAGCCGAGAGCTTGGCTGTGTCTTCCGACATGGTGTCTCCCTCCGGGTGATGAGCTCC 3353
QY 530 -----ProGlnArgGlyLeuGlnTyrLeuValGlyMetLeuArgGlnSerP 545
Db 3354 AGTACACACAGCTTACGGCTCCGGGGCCCGAGAGCTGCTGCGAGCTTCCGGAGTGTG 3413
QY 545 roAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyrL 565
Db 3414 CTATGTGGAGAGCTGCGGATGTGGAGCGCCACCATCTCAGTGTCCAGGACACACC 3473
QY 565 eu-----CysAlaValLeuGlnHisGlnGlyC 574
Db 3474 TGCAACACAGGATTCAGGATCAGCCTGAGATGATCCAGCTGTGCTGTGGACAGGAT 3533
QY 574 ySG:YLeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnL 593
Db 3534 GTTCTGTGAGACACTACAGCTGGAGCTTGAAGTGGTACTTGTGGACACAGCTCTT 3593
QY 594 LeuGlnAlaValLysArgAlaLysProAspLeuValLeuThrHisProAlaLeuAspGly 613
Db 3594 CTCAGGCTGTATCCATGAGGCTCA-----GCACTCCGACCCCGGCCCCCTGCTGT 3647
QY 614 HisProGlnProProLysGluLeu 621
Db 3648 TCAGGTTGGCCCTGCGCCGCTG 3671
```

RESULT 4

```
US-09-641-999-2
; Sequence 2, Application US/09641999
; Patent No. 6379894
; GENERAL INFORMATION:
; APPLICANT: MACH, BERNARD
; TITLE OF INVENTION: METHOD FOR SCREENING COMPOUNDS CAPABLE OF INHIBITING
; TITLE OF INVENTION: FIXING BETWEEN THE STAT. TRANSCRIPTION FACTOR AND THE
; TITLE OF INVENTION: USF1 TRANSCRIPTION FACTOR
; FILE REFERENCE: EGP 3.3-007CONT
; CURRENT APPLICATION NUMBER: US/09/641.999
; CURRENT FILING DATE: 2000-08-18
```

```
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4441
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-641-999-2

Alignment Scores:
Pred. No.: 4,768-22 Length: 4441
Score: 312.50 Matches: 196
Percent Similarity: 37.65% Conservative: 86
Rest Local Similarity: 26.17% Mismatches: 289
Query Match: 9.71% Indels: 181
DB: 4 Gaps: 29

US-10-029-347-2 (1-625) X US-09-641-999-2 (1-4441)

QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeuProAla 20
Db 1482 ATCTTGAAGAGACCTGACCGGCTTCTGCTCATCTTAGAGCCCTTCAGAGAGTGGAGCG 1541
QY 21 LeuGlyGlyProGluAlaAlaProCys---ThrAspProPheGluAlaAlaSerGlyAla 39
Db 1542 CAAGATGGCTTCTGCACAGACAGCTGGGACCGGACCGGCGAGCCCTGCTCCCTCCGG 1601
QY 40 ArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuValThr 59
Db 1602 GGGCTGCTGGCGGCTTTTCCAGAAAGAGTGTCTCCGAGGTGACACCTCTCTCTCTACA 1661
QY 60 ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla--- 78
Db 1662 GCCCGCCCCGG-----GCCCGCTGTCTCAGAGCCCTGAGCAAGGCCACCCCTATT 1715
QY 79 GluValArgGlyPheSerAspLysAspLysLysTyrPheTyrLysPhePheArgAsp 98
Db 1716 GAGCTGTCCGGCTTCTCCATGAGCAGGCGCCAGGATACGTATCGCTACTTTAGAGAGC 1775
QY 99 GluArgArgAlaGlu-----ArgAlaTyrArgPheValLysGluAsnGluThrLeu 115
Db 1776 TCAGGATGACAGAGCACCAGACAGAGCCCTGAGCTCTCCGGGACCCGGCCACTTCTT 1835
QY 116 PheAlaLeuCysPheValProPheValCysTrpIleValCysThrValLeuArgGlnGln 135
Db 1836 CTCAGTCACAGCCACAGCCCTACTTTGTCCGGGAGTGTGCGAGCTCTCAGAGGCCCTG 1895
QY 136 LeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyrLeu--- 154
Db 1896 CTGGAGCTTGGGAGGAC---GCCAAGCTGCCCTCCACGCTCAGCGGACTCTATGTGGC 1952
QY 155 LeuPheIleThrSerValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGly 174
Db 1953 CTGCTGGGCGTGTGAGCCCTCAGACGCCCTCC-----GGG 1988
QY 175 AspLeuArgAsnLeuCysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPhe 194
Db 1989 GCCTTGGCAGAGCTGGCCCAAGCTGGCTGGAG-----CTGGGCGGACAGATCAA--- 2039
QY 195 AlaGluLysGluLeuGlnLeuLeuArgGlySerLysValGlnThrLeuPheLeu 214
Db 2040 -----AGTACCCCTACAGGAGGACCATTCCTCCAGAGAGCTGAGGACCTGGGCGATG 2093
QY 215 SerLys-----LysGluLeuProGlyValLeuGluThrGluValThrTyr---Gln 230
Db 2094 GCCAAGGCTTAGTCCACACACCGCGGGCGGAGAGTCCGAGCTGGCTTCCCCAGC 2153
QY 231 PheIleAspGlnSerPheGlnGluPheLeu---AlaAlaLeuSerTyrLeuLeuGluAsp 249
Db 2154 TTCTCTCTCAATGCTTCTCTGGGGGCCCTGTGGCTGGCTCTCAGTGGCGAAATCAAGGAC 2213
QY 250 GlyGlyValProArgThrAlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGln 269
Db 2214 AAGGAGCTCCCGCAGTATCTTAGCA-----TTGACCCCAAGGAAGAGAGG 2258
```

QY	526	lnLeuProAspProGlnArgGlyLeuValGlyMetLeuArgGlnSerProA	546
DB	3268	GGCTGCGG-----GGGCC-CAGCAGCTCGCTCGCAGGCGTCC	3314
QY	546	laLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyrLeu-	565
DB	3315	ATGTGGAGAGCTGGCGATGTGGACGCCACCATCCCATTCAGTGTCCAGGAACACCTGC	3374
QY	566	-----CysAlaValLeuGlnHisGinglyCysG	575
DB	3375	AACAACAGGATTCACGGATCAGCCTGAGATGATCCAGCTGTGTCTCTGGACAGCATGTT	3434
QY	575	lyLeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuGlnGlu-Leu	594
DB	3435	CTCTGAGGACACTAACCACGCTGGACCTTGAACCTGGTACTGGTACCTGTGGACACAGCTCTTCTC	3494
QY	595	GlnAlaValIysArgAla-LysProAspLeuValIleThrHisProAlaLeuAspGlyHis	614
DB	3495	CAGCGCTGATCCCATGAGGCTCA-----GCATCTGGCACCGCGCCCTGCTGGTTC	3548
QY	615	ProGlnProProLysGluLeu 621	
DB	3549	GGGTGGCGCCCTGCCCGGCTG 3569	
RESULT 5			
US-08-295-502-:			
; Sequence 1, Application US/08295502			
; Patent No. 5672473			
; GENERAL INFORMATION:			
; APPLICANT: Glimcher, Laurie H.			
; APPLICANT: Zhou, Hong			
; APPLICANT: Douhan III, John			
; TITLE OF INVENTION: METHODS OF IDENTIFYING COMPOUNDS USEFUL			
; TITLE OF INVENTION: FOR TREATING AUTOIMMUNE DISEASES			
; NUMBER OF SEQUENCES: 1			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Fish & Richardson			
; STREET: 225 Franklin Street			
; CITY: Boston			
; STATE: MA			
; COUNTRY: USA			
; ZIP: 02110-2804			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.309			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/295,502			
; FILING DATE:			
; CLASSIFICATION: 436			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Freeman, John W.			
; REGISTRATION NUMBER: 29,066			
; REFERENCE/DOCKET NUMBER: 00246/188001			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 3393 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: DNA (genomic)			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: 1..3393			
US-08-295-502-1			
Alignment Scores:			
Pred. No.:		2,09e-21	3393
Score:		304.50	178
Percent Similarity:		37.54%	Conservative: 78
Best Local Similarity:		26.10%	Mismatches: 263
Query Match:		9.46%	Indels: 165

DB: : Gaps: 27

US-10-029-347-2 (1-625) x US-08-285-502-1 (1-3393)

QY : MetLeuA:adInProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLe:ProAla 20
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1459 A:CTTTGAAGAGACCTGCACCGCTTCTGCTCATCTCAGACGCC:TCGAGGAGACTGGAAAGCG :518

QY 21 LeuGlyGlyProGluA:AalaProCys---ThrAspProPheGluAlaAalaSerGlyAla 39
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1519 CAAGATGGC:TCTGTGCACACACACGTGGAGCCGACCGCGGAGCCCTGCTCCCTCGG 1578

QY 40 ArgValLeuGlyGlyLeuLeuSerLysAlaLeuProThrAlaLeuLeuValThr 59
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1579 GGGCTGCTGGCGGCC:TTTCCAGAGAAGCTGCTCGAGAGTTCCACCCCCTCTCTCACCA 1638

QY 60 ThrArgAlaAlaAalaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla--- 78
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1639 GCCCGGCCCGG-----GCCCGCTGCTCCAGAGCCTTGAGCAAGCGCGAGCCCCATT 1692

QY 79 GluValArgGlyPheSerArPlysAspLysLysTy:PheTyrLysPhePheArgAsp 99
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1693 GAGCTGTCGGCTCTCCATGGAGCAGCGCCGACGATACGTGATGCGCTACTTTGAGAGC 1752

QY 99 GluArgArgAlaGlu-----ArgAlaTyrArgPheValLysGluAsnGluThrLeu 115
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1753 TCAGGGATGACAGAGACCAGACAGAGCCTGAGCGCTCTCCGGAGCGGACACTCTT 1812

QY 116 PheAlaLeuCysPheValProPheValCysTrpIleValCysThrValLeuArgGlnGln 135
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1813 CTCACTCAGCCACACACCTACTTTTGTGGCGGAGTGTGCCAGCTCTCAGAGCGCCCTG 1872

QY 136 LeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyrLeu--- 154
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1873 CTGAGCTTGGGAGGAG--GCCAAGCTGCCCTCCAGGCTCAGCGAGCTCTATGTCCGC 1929

QY 155 LeuPheIleThrSerValLeuSerSerAlaproValAlaAspGlyProArgLeuGlnGly 174
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1930 CTGCTGGCGCTGCAGCGCTCGACAGCGCCGCC-----GGG 1965

QY 175 AspleuArgAsnLeu:CysArgLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPhe 194
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1966 GCCCTGCGAGCTGGCCAGCTGGCTGGGAG-----CTGGGCCGACAGACTCAA--- 2016

QY 195 AlaGlu-ySgLuLeuGluG:uGrLeuGluLeuArgGlySerLysVa:GlnThrLeuPheLeu 214
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 2017 -----AGTACCCTACAGGAGGACCGATTCCATCCGACAGCTGAGGAGCTGGCGCAT 2070

QY 215 SerLys-----LysGLeuProGlyValLeuGluThrGluValThrTyr---Gln 230
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 2071 GCCAAGGCTTAGTCCAACACCCACCGCGGCCGACAGTCCGAGCTGGGCTTCCCCAGC 2130

QY 231 Phe:IleAspGlnSerPheG:nGluPheLeu---AlaAlaLeuSerTyrLeuLeuGluAsp 249
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 2131 TTCCTCTGCAATCTCTCTGGGGCCCTGTGGCTGCTCTGAGTCGCGAAATCAGAGC 2190

QY 250 GlyGlyValProArgThrAlaAlaG:yGlyValGlyThrLeuLeuArgGlyAspAlaGln 269
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 2191 AAGGAGCTCCCGAGTACCTAGCA-----TTGACCCCAAGAGAGAAAGG 2235

QY 270 ProHisSerHis---LeuValLeuThrArgPheLeuPheGlyLeuLeuSerAlaGln 286
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 2236 CCTTATGACAACCTGGCTGGAGGGCTGCCACGCTTTCTGGCTGGGCTGATCTTCCAGCCT 2295

QY 289 ArgMetArgAspIleGluArgHisPheGlyCysMe:ValSerGluArgValLysGlnGlu 308
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 2296 CCCGCCGCTCCCTGGGAGTCTCTACTCGGCCATCGCGGCTGCTCGGTGGA-CAGGAA 2354

QY 309 AlaLeuArgn:pValGlnGlyGlnGlyCys--ProGlyValAlaAlaproGluValt 328
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 2355 GCAGA-HGGTCTTGGAGGTACTCTGAAGCGGCTCGACCGCGGG-----A 2398

QY 328 hrGluGlyA:ALysG:yLeuGluAspThrGluGluProGluGluGluGluGlyGluG 348

Db	2399	CACTCGGGCGCGGACGCTG		2418
Qy	348	luProAsnTyrProLeuGluLeuTyrCysLeuTyrGluThrGlnGluAspAlaPheV		368
Dd	2419	-----CITGAGCTGCTGACATGCGCCACGAGGCGGAGAGGCTGGAATT		2464
Qy	368	alArgGlnAlaLeuCysArgPhePro---GluLeuAlaLeuGlnArgValargPheCysA		387
Dd	2465	GGCAGCAGCTGGTACAGAGAGCTCCCGCGCGCCTCTCTTTCTGGGCACCGCGCTCAGC		2524
Qy	387	rgMetAspValAlaValLeuSerTyrCysValargCys		399
Dd	2525	CTCTGTATGCACATGTACTGGCAGGCGCTTGAGGGCGGCGGSCAAGACTTCTCCCTCG		2584
Qy	400	-----CysProIaagly---GlnAlaLeuArgLeuIleSerCys-		411
Dd	2585	ACCTCGCAGCACTGGCATTTCGCCCTCTGGATTGGGGAGCCTCGTGGGACTCAGCTGTG		2644
Qy	411	-----		411
Dd	2645	TCACCCTTTCAOGGCTGCTTGAGCGACACGGTGGCGCTGTGGGAGTCCCTCGCGGACG		2704
Qy	412	-----ArgLeuVaAlaAlaGlnGluLys		421
Dd	2705	ATGGGAGACCAAGCTACTTCAGGCAGCAGAGAGAATTCCACCATCGAGCCTTCCAAG		2764
Qy	421	ysLysSer-----LeuGlyLysArgLeuGlnAlaSerLeuGlyGlyG		435
Dd	2765	CCAAGTCCCTGAAGGATGTGGAAACCTGGGAAGCTTGTGCAGACTCAGAGGACGAGAA		2824
Qy	435	lySerSerGlnGlyThrThrLysGlnLeuProAla		447
Dd	2825	GTTCCTCGAAGACACAGCTGGGAGCTCCCTGCTGTTCGGGACCTAAGAACTGGAGT		2884
Qy	447	erLeuLeuHisProLeuPheGlnAlaMetThrAspPro		459
Dd	2885	TTGGCTGGGCGCTGTCTCAGGGCCCCCAGAGCTTCCCAAACCTGGTGGGATCTCACGG		2944
Qy	460	-----LeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspA		476
Dd	2945	CTTTTCTCTCGACATCTGCAGCTGGATGGCTGAGTGAGAACAGAGATCGGGACG		3004
Qy	476	laValCysA-gAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyL		496
Dd	3005	AGGTGTCTCGCAGCTCTCAGCCACCTTCCCCAGCTGAAGTCTTGAAACCCCTCAATC		3064
Qy	496	euLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpP		516
Dd	3065	TGTCACAGAACACATCCTGACCTGGTGCCTACAACTCGCGAGGCGCTG-----C		3118
Qy	516	roGlnCysArgValGlnThrValargValGln		526
Dd	3119	CTTCGCTCGTGATCCCTGCTCAGGCTAAGCTGTACAATACTGCATCTGCGACGTGG		3178
Qy	527	-----LeuProAsp		529
Dd	3179	GAGCCGAGAGCTTGGCTGTGTGCTTCCGACATGTGTCTCCCGGTGATGACGTCC		3238
Qy	530	-----ProGlnArgGlyLeuGlnTyrLeuValGlyMetLeuArgGlnSerP		545
Dd	3239	AGTACAACAAGTTACGGCTGCGGGGCCACGAGCTCGCTGCCACCTTCGGAGGTGTC		3298
Qy	545	roAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrYrL		565
Dd	3299	CTCATGTGGAGCGCTGGCATGTGGACGCCCACTCCCATTCAGTGTCCAGGAACACC		3358
Qy	565	eu 565		
Dd	3359	TG 3360		

RESULT 6
FCT-US95-10691-1

RESULT: 6
PCT-US95


```

; Sequence 1, Application PC/TUS95:10691
; GENERAL INFORMATION:
; APPLICANT: Glimcher, Laurie H.
; APPLICANT: Zhou, Hong
; APPLICANT: Dehan III, John
; TITLE OF INVENTION: METHODS OF IDENTIFYING COMPOUNDS
; TITLE OF INVENTION: USEFUL FOR TREATING AUTOIMMUNE DISEASES
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10691
; FILING DATE: 22 August 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/188031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3393
; PCT-US95-10691-1

```

```

Alignment Scores:
Pred. No.: 2,098-21 Length: 3393
Score: 304.50 Matches: 178
Percent Similarity: 37.54% Conservative: 78
Best Local Similarity: 26.10% Mismatches: 263
Query Match: 9.46% Indels: 165
DB: Gaps: 27

```

```
US-10-029-347-2 (1-625) x PCT-US95-10691-1 (1-3393)
```

```

QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeuProAla 20
DB 1459 ATCTTGAAGACCTGACCGGTTCTGCTATCTCTAGACGCTTCGAGGAGCTGGAGCG 1516
QY 21 LeuGlyGlyProGluAlaAlaProCys---ThraspProPheGluAlaAlaSerGlyAla 39
DB 1519 CAAGATGCTTCTCCACAGACGCTGCGGACCGGACCGGAGCGCTGCTCCCTCCGG 1578
QY 40 ArgValLeuGlyGlyLeuLeuSerCysAlaLeuLeuProThrAlaLeuLeuValThr 59
DB 1579 GGGCTGCTGGCGGCTTTCCAGAGAGAGCTGCTCCAGGTTGACACCTCTCTCACA 1638
QY 60 ThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCysAla--- 78
DB 1639 GCCCGGCCCGG-----GGCGGCTGCTCCAGACGCTGAGCAAGCGCGGCGCTATTT 1692
QY 79 GluAlaArgGlyPheSerAspLysAspLysLysTyrPheTyrLysPheArgAsp 98
DB 1693 GAGCTGTGCGGCTTCTCCATGAGCAGGCGCCAGGACATACGTGATGCGTACTTTGAGAGC 1752
QY 99 GluAlaArgAlaGlu-----ArgAlaTyrArgPheValLysGluAsnGluThrLeu 115
DB 1753 TCAGGATGACAGACGACCCAGACAGAGCCCTGAGCGCTCTCCGGGACCGGCCACTTCTT 1812

```

```

QY 116 PheAlaLeuCysPheValProPheValCysTrpIleValCysThrValLeuArgGlnGln 135
DB 1813 CTCAGTCCACAGCCACAGCCCTACTTTGTGCGCGGAGTGTCCAGACTCTCAGAGCCCTG 1872
QY 136 LeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyrLeu--- 154
DB 1873 CTGGAGCTTGGGAGGAC---GCCAAGCTGCCCTCCACGCTCACGGGACTCTATGTGGC 1929
QY 155 LeuPheIleThrSerValLeuSerSerAlaProValAlaAspGlyProArgLeuGlnGly 174
DB 1930 CTGCTGGGCGGTGCAGCCCTGCAGACCCCGCC-----GGG 1963
QY 175 AspLeuArgAsnLeuCysArgLeuAlaArgGluGlyValLeuGlyArgAlaGlnPhe 194
DB 1966 GCCCTGGCAGACTGGCCAGCTGGCTGGAG-----CTGGGCGGAGACATCAA--- 2016
QY 195 AlaGluLysGluLeuGluGlnLeuLeuArgGlySerLysValGlnThrLeuPheLeu 214
DB 2017 -----AGTACCTTACAGGAGGACCACTTCCATCCGAGACGTCAGGACCTGGGCGATG 2070
QY 215 SerLys-----LysGluLeuProGlyValLeuGluThrGluValThrTyr---Gln 230
DB 2072 GCCAAGGCTTAGTCCACACCCCGGCGCGAGAGTCCGAGCTGGCTTCCCGACG 2130
QY 231 PheIleAspGlnSerPheGlnGluPheLeu---AlaAlaLeuSerTyrLeuLeuGluAsp 243
DB 2132 TTCTCTCTGCAATGCTTCTTGGGGCCCTGTGGCTGGCTCTGAGTGGCGAAATCAAGGAC 2190
QY 250 GlyGlyValProArgThrAlaAlaGlyGlyValGlyThrLeuLeuArgGlyAspAlaGln 269
DB 2191 AAGGAGCTCCCGCAGTACTAGCA-----TTGACCCCAAGGAAGAAGG 2235
QY 270 ProHisSerHis---LeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGlu 288
DB 2236 CCCTATGACAACTGGCTGGAGGCGCTGCGCCACGCTTCTGGCTGGCTGATCTTCAGCT 2295
QY 259 ArgMetArgAspIleGluArgHisPheGlyCysMetValSerGluArgValGlyGlnGlu 308
DB 2296 CCGCGCGCTGCTGGAGCGCTACTCGGCGCATCGGCGGCTCGCTCGTGA-CAGGAA 2354
QY 309 AlaLeuArgTrpValGlnGlyGlnGlyCys---ProGlyValAlaProGluValTr 329
DB 2355 GCAGA-AGTGTCTGCGAGGTACCTGAAGCGGCTGCACCGGG-----A 2398
QY 328 TrGluGlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGlyGluG 348
DB 2399 CACTCGGCGCGGCGAGCTG----- 2418
QY 348 LuProAsnTyrProLeuGluLeuLeuTyrCysLeuTyrGluThrGlnGluAspAlaPheV 368
DB 2419 -----CTTGAGCTGTGCTGCTGCGCCACGAGCGCGAGGCTGGAATTT 2464
QY 368 aArgGluAlaLeuCysArgPhePro---GluLeuAlaLeuGlnArgValArgPheCysA 387
DB 2465 GGCAGACGTGTGACAGAGCTCCCGCGCGCTCTCTTTCTGGGACACCGCGCTCAGC 2524
QY 387 rgMetAspValAlaValLeuSerTyrCysValArgCys----- 399
DB 2525 CTCCTGATGCACATGCTACTGGGCAAGGCTTGGAGGCGCGGCGGCAAGACTTCTCCCTGG 2584
QY 400 -----CysProAlaGly---GlnAlaLeuArgLeuLeuSerCys- 411
DB 2585 ACCTCCGAGCAGCTGGCATTTGCGCTCTGATTTGGGAGGCTCTGTGGGAGCTAGCTGTG 2644
QY 411 ----- 411
DB 2645 TCACCGCTTTCAGGCGCTGCTTTCAGCGACACGCTGGCGCTGTGGAGTCCCTTCGCGGACG 2704
QY 412 -----ArgLeuValAlaAlaGlnGluLys-----LysL 421
DB 2705 ATGGGAGACCAAGACTACTTTCAGGCGAGGAGAGAGTTCACCATCGAGCTTTCAGAG 2764

```

```

QY 421 yLysSer-----LeuGlyLysArgLeuGlnAlaSerLeuGlyG:YG 435
D 421 yLysSer-----LeuGlyLysArgLeuGlnAlaSerLeuGlyG:YG 435
D 2765 CCAAGTCCCTGAGAGTGTGAGAGCTGCGAAAGCTGTGAGAGCTGAGAGAGAGAA 2824
QY 435 LysSerSerGlnGlyThrThrLysGlnLeuProAla-----S 447
D 435 LysSerSerGlnGlyThrThrLysGlnLeuProAla-----S 447
D 2825 GTCTCTCGAAGACACAGCTGCGAGCTCCCTGCTGTTCCGGAGCTTAAGAAACTGGAGT 2884
QY 447 eLeuLeuHisProLeuPheGlnAlaMetThrAspPro----- 459
D 447 eLeuLeuHisProLeuPheGlnAlaMetThrAspPro----- 459
D 2885 TTGGCTGGGGCTGCTGCTCAGAGCCCGAGGCTTTCCCAAACTGGTGGGATCTCAGCG 2944
QY 460 -----LeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspA 476
D 460 -----LeuCysHisLeuSerSerLeuThrLeuSerHisCysLysLeuProAspA 476
D 2945 CCTTTTCTCTCCCTGAGCATCTGGACCTGCGCTGAGTGAAGTGAAGTGAAGTGAAGT 3004
QY 476 laValCysArgAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyL 496
D 476 laValCysArgAspLeuSerGluAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyL 496
D 3005 AGGGTGCTCTCGAGCTCTCAGCCACCTTCCCGCAGCTGAAGTCTTGGAAACCTCAATC 3064
QY 496 eLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGlyGlyLeuAlaTTP 516
D 496 eLeuLeuHisAsnArgLeuSerGluAlaGlyLeuArgMetLeuSerGlyGlyLeuAlaTTP 516
D 3065 TGTCCAGAACAACTCACTGACCTGGGTGGTCAAACTCCGAGGCCCTG-----C 3128
QY 516 roGlnCysArgValGlnThrValArgValGln----- 526
D 516 roGlnCysArgValGlnThrValArgValGln----- 526
D 3119 CTTCGCTCGCTGCATCCCTGCTCAGCTTAAGCTTGTACATTAAGCTTGTACATTAAGCT 3178
QY 527 -----LeuProAsp----- 529
D 527 -----LeuProAsp----- 529
D 3179 GAGCCGAGAGCTTGGCTCGTGTCTTCCGAGCATGTGTCTCCGGGTGATGGAGCTCC 3238
QY 530 -----ProGlnArgGlyLeuGlnTyrrLeuValGlyMetLeuArgGlnSerP 545
D 530 -----ProGlnArgGlyLeuGlnTyrrLeuValGlyMetLeuArgGlnSerP 545
D 3239 AGTACAAAGTTCACGGTGTGGGGGCCAGAGCTCGCTCCAGCTTCCGAGGTGTC 3298
QY 545 roAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyrrL 565
D 545 roAlaLeuThrThrLeuAspLeuSerGlyCysGlnLeuProAlaProMetValThrTyrrL 565
D 3299 CTCATGTGGAGAGCTGGGGATGTGGACGCCCAACCATCCATTCAGTGTCCAGGAACACC 3358
QY 565 eu 565
D 3359 TG 3360

```

RESULT 7

```

US-09-099-041A-9
; Sequence 9, Application: US/09093041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/399,041A
; PRIOR FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-06-17
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2859
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-099-041A-9

```

Alignment Scores:

```

Pred. No.: 6,65e-20 Length: 2859
Score: 289.00 Matches: 172
Percent Similarity: 37.84% Conservative: 94
Best Local Similarity: 24.47% Mismatches: 261
Query Match: 159 Indels: 159
DB: 4 Gaps: 27

```

```

US-10-029-347-2 (1-625) x US-09-099-041A-9 (1-2859)
QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheAlaLeuAspGlyAlaAspGlyLeuProAla 20
D 1 MetLeuAlaGlnProGlnArgLeuLeuPheAlaLeuAspGlyAlaAspGlyLeuProAla 20
D 814 CTGCTGGCTTCCCGCCACGCTGGCCCTCTTCCATTCGATCGCTGACGAGCTGCATCG 873
QY 21 -----LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSer 37
D 21 -----LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSer 37
D 874 GACTTGGACCTGAGCCGCTGCTGACAGCTCTG-----CCCTGGAGACCTGCCAC 927
QY 38 GlyAlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeu 57
D 38 GlyAlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeu 57
D 928 CCCCTGCTCTTGTGGCAACCTGCTCAGTGGGAAGCTGCTCAAGGGGGTAGCAAGCTG 987
QY 58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCys 77
D 58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCys 77
D 988 CTCACAGCCCGCACA-----GGCATCAGAGTCCCGCCGAGTTCCTGCGGAAGAG 1038
QY 78 AlaGlnValArgGlyPheSerAspLysAspLysLysLysTyrrPheTyrrLysPhePheArg 97
D 78 AlaGlnValArgGlyPheSerAspLysAspLysLysLysTyrrPheTyrrLysPhePheArg 97
D 1039 GTGCTTCTCCGGGCTTCTCCCGCAGCCACTGCGCCCTATGCCAGGAGGATGTTCCCC 1098
QY 98 AspGluArgArgAlaGluArgAlaTyrrArgPheValLysGluAsnGluThrLeuPheAla 117
D 98 AspGluArgArgAlaGluArgAlaTyrrArgPheValLysGluAsnGluThrLeuPheAla 117
D 1099 GAGCGGGCCCTGCGAGGACCGCTGCTGAGCCAGCTGGAGGCCAACCCCAACCTCTGCAGC 1158
QY 118 LeuCysPheValProPheValCysTrrIleValCysThrVal-----LeuArgGln 134
D 118 LeuCysPheValProPheValCysTrrIleValCysThrVal-----LeuArgGln 134
D 1159 CTGTCCTCTGCGCCCTCTTCTGCTGATCATCTTCGGTGTCTCCAGCATCTCCGCT 1218
QY 135 GlnLeuLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyrrLeu 154
D 135 GlnLeuLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyrrLeu 154
D 1219 GCCTTGAAGGCTCACACAGCTGCCGAGTGCAGCATGACCTGCAGCATGCTCTCTC 1278
QY 155 LeuPheIleThrSerValLeuSer----- 162
D 155 LeuPheIleThrSerValLeuSer----- 162
D 1279 CTG---GTCACTGAGGTCCA-TCTGAACAGGATGCAGCCAGCAGCTGTGTGAGCGAA 1334
QY 163 ---SerAlaProValAlaAspGlyProArgLeuGlnGly-AspLeuArgLeuCysArg 181
D 163 ---SerAlaProValAlaAspGlyProArgLeuGlnGly-AspLeuArgLeuCysArg 181
D 1335 CACAGCGCCCATGTGAGAGACCTTCCACGCGCGCGGACACTCTGTGCTGTGTGGGCA 1394
QY 181 gLeuAlaArgGlyValLeuGlyArgAlaGlnPheAlaGlyLysGluLeuGluG 201
D 181 gLeuAlaArgGlyValLeuGlyArgAlaGlnPheAlaGlyLysGluLeuGluG 201
D 1395 GGTGCCCCACCGGGCATGGAGAAGACCTTCTTGTCTTCCAGCAGGAGGAGTGCAGGC 1454
QY 201 nLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysLeuPro-- 220
D 201 nLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysLeuPro-- 220
D 1455 CTCCGGGCTGCGAGGAGAGACATGACGCTGCGGCTTCTG-----CGGGCTTGGCGGA 1508
QY 221 ---GlyValLeuGluThrGluValThrTyrrGlnPheIleAspGlnSerPheGlnGluPh 239
D 221 ---GlyValLeuGluThrGluValThrTyrrGlnPheIleAspGlnSerPheGlnGluPh 239
D 1509 GCTGGGCCCGGGGTGACGAGCATCTATAGTTTTCACCTCACCCTCAGGCTT 1568
QY 239 eLeuAlaAlaLeuSerTyrrLeuLeuGluAsp----- 249
D 239 eLeuAlaAlaLeuSerTyrrLeuLeuGluAsp----- 249
D 1569 CTTCACAGCTTCTCTCTGCTGCTGGACACAGAGTGGGCTGCTGAGTTCACCTCAGGCTT 1628
QY 250 -----GlyGlyValProArgThrAlaAla----- 257
D 250 -----GlyGlyValProArgThrAlaAla----- 257
D 1629 CTTCAGAGTGGATGCCCTTCCGGGGGAGGAGCCACCTCTCTCTATCTCCCTTCT 1688
QY 258 -----GlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisSe 272
D 258 -----GlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisSe 272
D 1689 CCGTTCAGTGCCTGCGGGGAGTGTGCGGGGAGAGAGCTCTTCAAGAACAAAGGA 1748
QY 272 rHisLeuValLeuThrThrArgPhePheGlyLeuLeuSerAlaGluArgMetArgAs 292
D 272 rHisLeuValLeuThrThrArgPhePheGlyLeuLeuSerAlaGluArgMetArgAs 292
D 1749 TCATCTCCAGTTTCAACCACTCTTCTGCTGGGGCTGTGTCTCCAAAGCCCAACAGAAA-- 1806
QY 292 pIleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTr 312
D 292 pIleGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTr 312
D 1807 -CTCTGCGGCATCTGTGTGCGCGGAGCCCTCAGGAGAAAGCGCAAGGCCCTCTG---TG 1862

```

```

CY 312 p-----ValGlnGlyGlnGlyGlnGlyCysProGlyValAlaAProG1 326
Db 1863 GGCACACCTGTTTCCAGCTTGGGGGCTACCTGAAAGAGCCTGCCCGGCTTCAAGTCCA 1922
CY 326 uValThrGluGlyAlaLysGlyLeuGluAlaSerThrGluGluProGluGlyGluGluGluG1 346
Db 1923 AAGCTTCAACGAGTGCAGGCCATG----- 1947
CY 346 yGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyrGluThrGlnGluAsp-- 365
Db 1948 -----CCCACGCTC---ATCGGATGCTGCCTGCATCTACGACACACAGAGCCAGAA 1997
CY 366 -----AlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuGluAlaGva 383
Db 1998 GTGGGGCAGCTGGCGCCAGGGCATCTCGCC---AACTACCTCAAGCTGACCTACTG 2054
CY 383 lArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCysProAlaG1 403
Db 2055 CAACGCTGCTCGCGGAGCTGCAGGCGCTCTCTGCTGCTCATCACTTCCCGAAGCG 2114
CY 403 yGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluLysLysCys 423
Db 2115 GCTGGCCCTAGACTAGACACAAACAATCTC----- 2145
CY 423 rLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGlyThrThrLysG1 443
Db 2146 -----AAGACTACGGGCTGGGGAG-- 2166
CY 443 nLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLe 463
Db 2167 -----CTGACGCCCTGCTTCAGCGCTCACT----- 2193
CY 463 uSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgLeuSerG1 483
Db 2194 -----GTTCTCAGACTCAGCTAAACAGATCACTGACCGTGGGGTAAAGGCTTAGCGA 2249
CY 483 uAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAspArgLeuSerG1 503
Db 2250 AGAGCTGACCAAAATACAAAATCTGACCTATTGSGGTTATACAAACACCATCAGCAGA 2309
CY 503 uAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrVa 523
Db 2310 TGTGGGAGCCAGGTACGTACCAAAAATCTCTG-----GATGAATGCAAGGCTCAGC-- 2361
CY 523 lArgValGlnLeuProAsp-----ProGlnArgGlyLeuGlyLeu----- 537
Db 2362 -CATCTTAACTGGGAAAAAACAATAACAGTGAAGGAGGGAAGTATCTCGCCCTGGC 2420
CY 538 -----ValGlyMet----- 540
Db 2421 TGTGAAGAACAGCAATCAATCTCTGAGGTGGGATGTGGGGCAATCAAGTTGGGGATGA 2480
CY 541 -----LeuArgGlnSerProAlaLeuThrThrLeuAspIle 552
Db 2481 AGAGCAAAAGCCCTTCGACAGAGCTCTGGCGGAACCAACCCGAGCTTACCAACCTGAGTC 2540
CY 552 uSerGlyCysGlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuGlnHisG1 572
Db 2541 TGGCTCCACGGCATCTCCACAGAGGAGGAAGAGCCCTTGGAGGGCTCTGCAG---CA 2597
CY 572 rGlyCysGlyLeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuG1 592
Db 2598 GAACACGCTCTAGAAATACTGTGGCTGACCCAAAATGAATCAACGATGAAGTGGCAGA 2657
CY 592 rGluLeuGlnAlaValLysArgAlaLysProAspLeu-----ValLeuThrHisProAl 610
Db 2658 GAGTTTGGCAGAAATGTT-GAAAGTCAACACAGACGTTAAAGCATTTTATGCTTATCCAGA 2716
CY 610 aLeuAsp 612
Db 2717 ATCAGAT 2723

```

RESULT 8

```

US-09-245-281-9
; Sequence 9, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 07334/118001
; CURRENT APPLICATION NUMBER: US/09/245,281
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: US 09/207,359
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 09/099,041
; EARLIER FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: US 09/019,942
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2859
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-281-9

```

```

Alignment Scores:
Pcted. No.: 6,65e-20 Length: 2859
Score: 289.00 Matches: 172
Percent Similarity: 37.84% Conservative: 94
Best Local Similarity: 24.47% Mismatches: 281
Query Match: 8.98% Indels: 158
DB: 4 Gaps: 27

```

US-10-029-347-2 (1-625) x US-09-245-281-9 (1-2859)

```

QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheLeuAspGlyAlaAspGluLeuProAla 20
Db 814 CTGCTGGCTTCCCGCCAGCTGGCCCTCTTCACCTTCGATGCCCTGGACGAGCTGCACTCG 873
QY 21 -----LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSer 37
Db 874 GACTTGGACCTGAGCCCGCTGCTGACAGCTCTCTGC-----CCCTGGGAGCTGGCCAC 927
QY 38 GlyAlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuLeu 57
Db 928 CCCCTGCTCTGTGGCCCAACCTGCTCAGTGGAGAGCTGCTCAAGGGGTAGCAAGCTG 987
QY 58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCys 77
Db 988 CTCACAGCCCGCAC-----GGCATCGAGGTCCCGGCCAGTTCCTGCGGAAGAAG 1038
QY 78 AlaGluValArgGlyPheSerAspLysAspLysLysLysLysLysPheTyrLysPhePheArg 97
Db 1039 GTGCTTCTCGGGCTTCTCCCGAGCCACCTCGCGCCCTATGCCAGGAGGATGTTCCCC 1098
QY 98 AspGluArgArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAla 117
Db 1099 GAGGGGCGCTCGACAGCCGCTGCTGAGCAGCTGGAGGCCAACCCCAACCTCTGCAGC 1158
QY 118 LeuCysPheValProPheValCysTrpIleValCysThrVal-----LeuArgGln 134
Db 1159 CTGTGCTCTGTGCCCTCTTCTGCTGATCATCTTCGGTGTCTTCCAGCAGCTTCCCTGT 1218
QY 135 GlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyrLeu 154
Db 1219 GCCTTTGAAGGCTCACACAGCTGCCGAGCTGCAGATGACCTGACAGATGCTTCTTC 1278
QY 155 LeuPheLeuThrSerValLeuSer----- 162
Db 1279 CTG---GTCACCTGAGGTCCA-TCTGAACAGGATGACGCCAGCAGCTGGTGCAGCGGA 1334
QY 163 ---SerAlaProValAlaAspGlyProArgLeuGlnGly-AspLeuArgAsnLeuCysAr 181
Db 163 ---SerAlaProValAlaAspGlyProArgLeuGlnGly-AspLeuArgAsnLeuCysAr 181

```

```
Db 1335 CACACGAGCCAGTGGAGAGCTCCACCGCGCGGCGGACACACTCTGTGTCTCGCTGGCGGCA 1394
Qy 181 gLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 201
Db 1395 GGTGGCCACCGCGGCATGGAGAGAGCCCTCTGTCTTCAACCACGAGGAGGTGCAGGC 1454
Qy 201 nLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro-- 220
Db 1455 CTCGGGCTCGAGGAGAGACATGGACGCTGGCTTCTCTG-----CGGGCTTGGCCGA 1508
Qy 221 ----GlyValLeuGluThrGluValThrTyThrGlnPheIleAspGlnSerPheGlnGluPh 239
Db 1509 GCTGGCCCGCGGGGTGACACGAGCTCTATGAGTCTTTTCCACCTCACCCCTCCAGGCTT 1568
Qy 239 eLeuAlaAlaLeuSerTyThrLeuLeuGluAsp----- 249
Db 1569 CTTTACAGCCTTCTTCTCGTGTGGAGACAGAGGTGGGCACTCAGGAGCTGTCTCAGGTT 1628
Qy 250 -----GlyGlyValProArgThrAlaAla----- 257
Db 1629 CTTCAGAGTGGATGCCCTCGCGGGGCGAGCGACCTCTCTCTCTATCTCTCTCTCTCT 1688
Qy 258 -----GlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisLe 272
Db 1689 CCCGTTCCAGTGCCTGAGGACGCTGTGTCCGGCGCGGAGACCTCTTCAAGAACACAGGA 1748
Qy 272 rHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAs 232
Db 1749 TCACCTCCAGTTCCACCAACCTCTCTCTGTGGGGCTGTGTCCAAAGCCCAACAGAAA-- 1806
Qy 292 pLeuGluArgHisPheGlyCysMetValSerGluArgValLysGlnGlnAlaLeuArgTr 312
Db 1807 -CTCTCGGGCATCTGTGTCCCGCGGAGCCCTCAGGAGAAAGGCGAGGCCCTG---TG 1862
Qy 312 p-----ValGlnGlyGlnGlyCysProGlyValAlaAProG 326
Db 1863 GGCACACCTGTTTCCAGCTTCCGGGGCTACCTGAAGAGCGCTGCCCGCGTTCAGGTCGA 1922
Qy 326 uValThrGluGlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGlu 346
Db 1923 AAGTTCAACAGGTGCAGGCCATG----- 1947
Qy 346 yGluGluProAsnTyProLeuGluLeuLeuTyThrCysLeuTyThrGlnGlnGluAsp-- 365
Db 1948 -----CCCAGTCT---ATCTGGAGCTCGGTGTGATCTACGAGACACAGAGCCAGAA 1997
Qy 366 -----AlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgva 383
Db 1998 GGTGGGCGAGCTGGCGGCGAGGGCATCTGGCC---AACTACCTCAAGCTGACCTACTG 2054
Qy 383 lArgPheCysArgMetAspValAlaValLeuSerTyThrCysValArgCysCysProAlaG 403
Db 2055 CAAGCGCTGTCTGGCGGAGCTGAGCGCCCTCTCTCTGCTGCTGATCATCTTCCCAAGCG 2114
Qy 403 yGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluLysLysLysLe 423
Db 2115 GCTGGCCCTAGACCTAGACACACACATCTC----- 2145
Qy 423 rLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGlyThrThrLysG 443
Db 2146 -----AACGACTACGGCGTGGCGGAG-- 2166
Qy 443 nLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLe 463
Db 2167 -----CTGCGAGCCCTGCTTCAAGCGCCTCACT----- 2193
Qy 463 userSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerG 483
Db 2194 ----GTTCTCAGACTCAGCGTAAACACGATCCTACGCGTGGGTAAAGGTGTCTAAGCA 2249
Qy 483 uAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerG 503
Db 2250 AGAGCTGACCAAAATACAAAATTTGTGACCTATTGTGGTTTATACAAACACACAGATCACC 2309
```

```
Qy 503 uAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrVa 523
Db 2310 TGTCCGAGCCAGGTACGTACGTCACCAAAATCTGT-----GATGAATGCAAAAGGCTCAGC-- 2361
Qy 523 lArgValGlnLeuProAsp-----ProGlnArgGlyLeuGlnTyThrLeu----- 537
Db 2362 -CATCTTAACTGGGAAAAAACAAAATAACAAGTGAAGGAGGAAAGTATCTCGCCCTGGC 2420
Qy 538 -----ValGlyMet----- 540
Db 2421 TGTGAAGACACGAAATCAATCTCTGAGTTGGGATGTGGGCAATCAAGTTGGGATGA 2480
Qy 541 -----LeuArgGlnSerProAlaLeuThrThrLeuAspLe 552
Db 2481 AGAGCAAAAGCCTTCGAGAGGCTCTCGGGAACACCCAGCTTGACCCCTTAGTCT 2540
Qy 552 uSerGlyCysGlnLeuProAlaProMetValThrTyThrLeuCysAlaValLeuGlnHisG 572
Db 2541 TGGTCCACAGGCATCTCCACAGAGGAGGAAAGAGCCTTGGGAGGGCCCTGCAG--CA 2597
Qy 572 nGlyCysGlyLeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuG 592
Db 2598 GAACACGCTCTCTAGAAATACTGTGGCTGACCCCAAAATGAACCTCAACGATGAAGTGCAG 2657
Qy 592 nGluLeuGlnAlaValLysArgAlaLysProAspLeu-----ValIleThrHisProAl 610
Db 2658 GAGTTGGCAGAAATGTT-GAAAGTCAACACGACGTTAAAGCATTTATGGCTTATCCAG 2716
Qy 610 aLeuAsp 612
Db 2717 ACAGAT 2723

RESULT 9
US-09-207-359B-9
; Sequence 9, Application: US/09207359B
; Patent No. 6469140
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-112001
; CURRENT APPLICATION NUMBER: US/09/207,359B
; CURRENT FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2859
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-207-359B-9

Alignment Scores:
Pred. No.: 6,65E-20 Length: 2859
Score: 289.00 Matches: 172
Percent Similarity: 37.84% Conservative: 94
Best Local Similarity: 24.47% Mismatches: 281
Query Match: 8.98% Indels: 158
DB: 4 Gaps: 27

US-10-029-347-2 (1-625) x US-09-207-359B-9 (1-2859)
Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
Db 814 CTGCTGGCTTCCCCACCGCTGCTTCTTCACTTCGATGGCTGGAGGCTGCACTCG 873
Qy 21 -----LeuGlyGlyProGluAlaAlaProCysThrAspPheGluAlaAlaSer 37
Db 874 GACTTGGAGCTGAGCGCGCTGCTGACAGCTCTCTG-----CCCTGGGAGCTTGGCCAC 927
```

Qy 38 GlyAlaArgValLeuG:yglyLeuLeuSerlyslsAlaLeuLeuProThrAlaLeuLeuLeu 57
Db 928 CCCCTGGTCTGGTGGCAACTCTCTCAGTGGGAAGCTGCTCAAGGGGGCTAGCAAGCTG 987
Qy 58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCys 77
Db 988 CTCACAGCCCGCAC-----GGCATCGAGGTCCTCCGCGCAGTTCCTCGCGAAGAG 1038
Qy 78 AlaGluValArgGlyPheSerAspLyAspLylyslsTyrPheTyrLylyslsPhePheArg 97
Db 1039 GTGCTCTCCGGGCTCTCTCCAGCACCTCGCCCTATGCCAGGAGGATGTTCCCC 1098
Qy 98 AspGluArgArgAlaGluArgAlaTyrArgPheVallyslsGluLysGlnThrLeuPheAla 117
Db 1099 GAGCGGGCCCTGCGAGGCGCTCTCGAGCGAGCTGGAGGCCAACCCCACTCTCGAGC 1159
Qy 118 LeuCysPheValProPheValCysTrpIleValCysThrVal-----LeuArgGln 134
Db 1159 CTGTGCTGTGGCCCTCTTCTGTGTGATCATCTTCGGTGCTTCCAGCACTTCGGTGT 1218
Qy 135 GlnLeuGluLeuGlyArgAspLeuSerArgThrSerLylyslsThrThrThrSerValTyrLeu 154
Db 1219 GCCTTTGAGGCTCAACACAGCTCCCGACCTGCAGATGACCTGACAGATGTCTCTCTC 1278
Qy 155 LeuPheIleThrSerValLeuSer-----162
Db 1279 CTG---GTACTGTAGGTCCA-TCTGACAGATGACGCCAGCGCTGTGTGAGCGGAA 1334
Qy 163 ---SerAlaProValAlaAspGlyProArgLeuGlnGly-AspLeuArgAsnLeuCysAr 18:
Db 1335 CACAGCGAGCCAGTGGAGACCTCCAGCGCGCGGACACCTCTGTGCTCGTGGGCA 1394
Qy 181 GluAlaArgGlnGlyValLeuGlyArgArgAlaGlnPheAlaGlnGlyLeuGluG: 20:
Db 1395 GGTGCCACCGGGGCGATGGAGAGAGCTCTTGTCTTCACCGAGGAGGTGCGAGC 1454
Qy 201 nLeuGluLeuArgGlySerLylyslsValGlnThrLeuPheLeuSerLylyslsGluLeuPro-- 220
Db 1455 CTCGGGCTGACGAGAGACATGACAGCTGGGCTTCTG-----CGGGTGTTCGCGA 1509
Qy 221 ---GlyValLeuGluThrG:uValThrTyrGlnPheLeuAspGlnSerPheGlnG:uph 239
Db 1509 GCTGGGCGCGGGGTGACGAGCATCTCTATGASTTTTCACCTCACCCTCCAGGCTT 1568
Qy 239 eLeuAlaAlaLeuSerTyrLeuLeuGluAsp-----249
Db 1569 CTTTACAGCTTCTTCTCTCGTCTGGAGCAGAGGGTGGGCACTCAGGAGTGTCTCAGGTT 1628
Qy 250 -----GlyGlyValProArgThrAlaAla-----257
Db 1629 CTTCCAGAGTGGATGCCCTTGGGGGCGAGCGACCGTCTCTCTATCTCTCCCTTCT 1688
Qy 258 -----GlyG:lyalG:lyThrLeuLeuArgGlyAspAlaGlnProHis 272
Db 1689 CCGGTTCCAGTGTCTGCGAGGCGAGTGGTGGCGCGCGGAGACCTCTTCAAGAACAGGA 1748
Qy 272 rHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAs 292
Db 1749 TCAGTCCAGTTCACCACTCTTCTCTGTGGGCTGTGTCCAAAGCCAAACAGAAA-- 1806
Qy 292 pIleGluArgHisPheGlyCysMetValSerGluArgVallyslsGlnGluAlaLeuArgTr 312
Db 1807 -CTCTCGGCGATCTGGTGGCGCGGAGCCCTGAGGAGAAAGCGCAAGGCCCTG---TG 1862
Qy 312 p-----ValGlnGlyGlnGlyCysProGlyValAlaProG 326
Db 1863 GGCACACCTGTTTCCAGCTGCGGGGTACTGGAAGAGCCGTGCCCCGCTTCAGTCTGA 1922
Qy 326 uValThrGluGlyAlaLylyslsGlyLeuGluAspThrGluGluGluGluG:uG 346
Db 1923 AAGCTTCAACAGGTGCGAGGCCATG-----1947

Qy 346 yGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyrGlnThrGlnGluAsp-- 365
Db 1948 -----CCCAGCTTC---ATCTGATGCTGGCTGTCATCTACAGACACAGAGCCAGAA 1997
Qy 366 -----AlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgVa 383
Db 1998 GGTGGGGCAGCTGGGGCCAGGGGCATCTGGCC---AATCTCTCAAGCTGACCTACTG 2054
Qy 383 lArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCysProAlaG 403
Db 2055 CAAGCCCTGCTCGGCGCAGCTGCAGGCCCTCTCTCTGCTGTCATCTCTCCCAAGCG 2114
Qy 403 yGluAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGlnLylyslsLyse 423
Db 2115 GCTGGCCCTAGACCTAGACACAAATCTC-----2145
Qy 423 rLeuGlyLylyslsArgLeuGlnAlaSerLeuGlyGlyGlyserSerGlnGlyThrThrlysl 443
Db 2146 -----AACGACTACGGCGTGGCGAG-- 2166
Qy 443 nLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLe 463
Db 2167 -----CTGAGCCCTGCTTACGCCGCTCACT-----2193
Qy 463 uSerSerLeuThrLeuSerHisCysLylyslsLeuProAspAlaValCysArgAspLeuSerG 483
Db 2194 ---GTTCTCAGACTCAGCGTAACACAGATCAGCTGCGTGGGTAAAGTGTCTAAGCGA 2249
Qy 483 uAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerG 503
Db 2250 AGAGCTGACCAATAACAAATTTGAGCTATTGCGTTTATACAAACACAGATCACCGA 2309
Qy 503 uAlaGlyLeuArgMetLeuSerGlyLeuAlaTrpProGlnCysArgValG:nThrVa 523
Db 2310 TGTCGAGCCAGGTACGTCCAAATCTCTG-----GATGAATGCAAGGCCTCAG-- 2361
Qy 523 lArgValGlnLeuProAsp-----ProGlnArgGlyLeuGlnTyrLeu-- 537
Db 2362 -CATCTTAACTGGGAAAAACAAATAACAAGTGAAGGAGGAGTATCTCGCCCTGGC 2420
Qy 538 -----ValGlyMet-----540
Db 2421 TGTGAAGAACACAAATCAATCTCTGAGGTTGGATGTGGGGCAATCAAGTGGGGATGA 2480
Qy 541 -----LeuArgGlnSerProAlaLeuThrThrLeuAspLe 552
Db 2481 AGGACAAAGCCTTCGCGAGGCTCTCGGAACACCCAGCTTGACCACTGAGTCT 2540
Qy 552 uSerGlyCysGlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuGlnHisG 572
Db 2541 TCGGTCCAACGCGATCTCCACAGAGGAGGAGGCTTCGAGGGGCTTGCAG--CA 2597
Qy 572 nClyCysGlyLeuGlnThrLeuSerLeuAlaSerValGluLeuSerG:uGlnSerLeuG 592
Db 2598 GAACAGCTCTCTAGAAATACTGTGGCTGACCAATAATGAACTACAGATGATGAGTGGCAGA 2657
Qy 592 nGluLeuGlnAlaVallyslsArgAlaLylyslsProAspLeu-----ValIleThrHisProAl 610
Db 2658 GAGTTTGGCAGAAATGTT-GAAGTCAACACAGAGCTTAAAGCATTTATGGCTTATCCAGA 2716
Qy 610 aLeuAsp 612
Db 2717 ATCAGAT 2723

RESULT 10

US-09-340-620A-9
; Sequence 9, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A

```

; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-620A-9

Alignment Scores:
Pred. No.: 6,65e-20 Length: 2859
Score: 289.00 Matches: 172
Percent Similarity: 37.84% Conservative: 94
Best Local Similarity: 24.47% Mismatches: 281
Query Match: 8.95% Indels: 158
DB: 4 Gaps: 27

US-10-029-347-2 (1-625): x US-09-340-620A-9 (1-2859)
QY 1 MetLeuAlaGlnProGluArgLeuLeuPheLeuLeuAspGlyAlaAspGluLeuProhla 20
DB 8.4 CTGTGGCGCTTCCCGCCAGCTGGGCCCTCTTACCTTCGATCGCTGGAGAGCTGCACCTG 873
QY 21 -----LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSer 37
DB 874 GACTTGGACCTGAGCGCGTGGCTGACAGCTCTCTG-----CCCTGGAGCGCTGCCAC 927
QY 38 GlyAlaA-gValLeuGlyGlyLeuLeuSerTysAlaLeuLeuProThrAlaLeuLeuLeu 57
DB 928 CCCCTGGTCTTCTGCTGGCCACCTGCTCAGTGGGAAGCTGCTCAAGGGGGCT-AGCAAGCTG 987
QY 58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCys 77
DB 988 CTCACAGCGCCGACACA-----GGCATCGAGGTCCCGCGGCAGTTCTCTCGGAAGAG 1036
QY 78 AlaGluValArgGlyPheSerAspLysLysLysLysLysLysLysLysLysLysLysLys 97
DB 1039 GTGCTTCTCCGGGGTCTTCTCCCGAGCCACCTGCGCGCTATGCCAGAGATGTTCGCC 1098
QY 98 AspGluArgArgAlaGluArgAlaTyArgPheValLysGluAsnGluThrLeuPheAla 117
DB 1099 GAGCGGGCGCTCCAGACCGCTGCTGAGCCAGCTGGAGGCCAACCCCAACCTCTGCAGC 1158
QY 118 LeuCysPheValProPheValCysTrpIleValCysThrVal-----LeuArgGln 134
DB 1159 CTGTGCTGTGCGCCCTCTTCTGTGATCATCTTCCCGTGCTTCCAGCAGCTTCGCTGCT 1218
QY 135 GlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyLeu 154
DB 1219 GCCTTTGAAGGCTCACCAGCTGCCGAGCTCCAGATGATCATCTTCCCGTGCTTCCAGCAGCTTCGCTGCT 1279
QY 155 LeuPheLeuThrSerValLeuSer----- 162
DB 1279 CTG---GTCACTGAGTCCA-TCTGAACAGGATGACGCCCGCCAGCTGCTGTCAGCGGAA 1334
QY 163 ---SerAlaProValAlaAspGlyProArgLeuGlnGly-AspLeuArgAsnLeuCysAr 181
DB 1335 CACACGACGACGAGTGGAGACCTCCACCGCGCGGAGACATCTGTGCTGCTGGGCA 1394
QY 181 sLeuAlaAspGluGlyValLeuGlyArgArgAlaGlnPheA-aGluLysGluLeuGluG 201
DB 1395 GTTGGCCACCGGGGATGGAGAAGAGCTCTTTCTTTCACCCCGAGGAGGTGACGC 1454
QY 201 nLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro-- 220

```

```

DB 1455 CTCGGGGCTGCAGGAGAGACATGCAGCTGGGCTTCCTG-----CGGGCTTTGCCGA 1508
QY 221 ----GlyValLeuGluThrGluValThrTyGlnPheIleAspGlnSerPheGlnGluPh 239
DB 1509 GCTGGGCCCCGGGGTGACACGAGTCTTATGAGTTTTTCCACCTCACCCCTCCAGGCTT 1568
QY 239 eLeuAlaAlaLeuSerTyLeuLeuGluAsp----- 249
DB 1569 CTTACAGCTTCTTCTGCTGCTGACACAGGGTGGGCACTCAGGAGCTGCTCAGGTT 1628
QY 250 -----GlyGlyValProArgThrAlaAla----- 257
DB 1629 CTTCCAGGAGTGATCCCCCTGCGGGGCGAGCGACACGCTGCTGCTATCTCTCCCTTCT 1688
QY 258 -----GlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHisSe 272
DB 1689 CCGTTCAGTGCCTCAGGGCAGTGGTCCGGCGGGAGAGCCTTTCAAGAACCAAGGA 1748
QY 272 rHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAs 292
DB 1749 TCACCTCCAGCTTCACCAACCTCTCTGCGGGCTGTGTCCAAAGCCAAACAGAAA-- 1806
QY 292 pIleG-uArgHisPheGlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTr 3:2
DB 1807 -CTCTGGGCACTGTGTGCCCGCGCAGCCCTGAGGAGAAAGCGAAGCCCTG---TG 1862
QY 312 p-----ValGlnGlyGlnGlyG-nglyCysProGlyValAlaProG 326
DB 1863 GGCACACCTGTTTCAGCGCTCGGGGCTACCTGAAGAGCCTGCCCGGCTTCAGGTGCA 1922
QY 326 uValThrGluGlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGluG 346
DB 1923 AAGCTTCACACAGGTGCAGGCCATG----- 1947
QY 346 yG-uGluProAsnTyProLeuGluLeuLeuTyCysLeuTyGluThrGlnGluAsp-- 365
DB 1948 -----CCCACGCTTC---ATCTGGATGTCTGCTGCTGCTGCATCCTCCCAAGCG 1997
QY 366 -----AlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgVa 383
DB 1998 GTGGGGGACGCTGGCGCCAGGGGCTCTGCGCC---AACTACCTCAAGCTGACCTACTG 2054
QY 383 lArgPheCysArgMetAspValAlaValLeuSerTyCysValArgCysCysProAlaG 403
DB 2055 CAAGCGCTGTGCGCCGACGTGCAGCGCTCTCTCTGCTGCTGCATCCTCCCAAGCG 2114
QY 403 yGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluLysLysLysSe 423
DB 2115 GCTGGCCCTAGACCTAGACAAACAATCTC----- 2145
QY 423 rLeuGlyLysArgLeuGlnAlaSerLeuGlyGlySerSerGlnGlyThrThrLysG 443
DB 2146 -----AACCACTACGGCTGGGGAG-- 2166
QY 443 rLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLe 463
DB 2167 -----CTGCAGCGCTGCTTCAGCGCTCACT-- 2193
QY 463 uSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerG 483
DB 2194 ----GTCTCAGACTCAGCGTAACAGACACTCAGTGGGTAAAGGTGCTAAAGCGA 2249
QY 483 uAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerG 503
DB 2250 AGAGCTGACCAATAACAAATTGTGACCTATTGGGTATTATACAAACACCATCACCAG 2309
QY 503 uAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrVa 523
DB 2310 TGTGGAGCCAGGTACGTCCACCAATCTCTG-----GATGAATGCAAAAGCCCTCAG-- 2361
QY 523 lArgValGlnLeuProAsp-----ProGlnArgGlyLeuGlnTyLeu----- 537
DB 2362 -CATCTTAAACTGGAAAAAACAATAACAAGTGAAGGAGGAAGTATCTCCCTGCGC 2420

```

```

Qy 538 -----ValGlyMe----- 540
Db 2421 TGTGAAGAACAGCAATCAATCTCTGAGGTGGGATCTGGGGCAATCAAGTTGGGGATGA 2480
Cy 541 -----LeuArgInSerProAlaLeuThrThrLeuAspLe 552
Db 2481 AGGAGCAAAAGCCTTCGAGAGGCTCTGGGGAACCAACCCAGCTTGACCACTGAGTCT 2540
Qy 552 uSerGlyCysGlnLeuProAlaProMetValThrTyLeuCysAlaValLeuGlnHisG1 572
Db 2541 TGCCTCCAACGGCATCTCCAAGAGGAGGAAAGAGCCTTCGAGGGCCCTGCAG---CA 2597
Qy 572 nGlyCysGlyLeuGlnThrLeuSerLeuAlaSerValGLeuLeuSerGluGlnSerLeuG1 592
Db 2598 GAACACGCTCTTGAATACTGTGGTACCCCAAAATGAATCAACGATGAAGTGGCAGA 2657
Qy 592 nGluLeuGlnAlaVallysArgAlaLysProAspLeu-----ValIleThrHisProAl 610
Db 2658 GAGTTTGGCAAAATGTT-GAAGTCAACCAAGACGTTAAAGCATTTATGGCTTATCCAGA 2716
Qy 610 aLeuAsp 612
Db 2717 ATCAGAT 2723

RESULT 11
US-09-099-041A-7
; Sequence 7, Application JS/0909941A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPL.CANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 97334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/619,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (245)...(3103)
US-09-099-041A-7

Alignment Scores:
Pred. No.: 8.8e-20 Length: 3392
Score: 289.00 Matches: 172
Percent Similarity: 37.84% Conservative: 94
Best Local Similarity: 24.47% Mismatches: 281
Query Match: 8.98% Indels: 158
Gaps: 27

US-10-029-347-2 (-625) x US-09-099-041A-7 (1-3382)
Qy 1 XetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
Db 1058 CTGTGTGGCTTCCGCCACGTGGCCCTCTTCACTTCGATGGCCTGGAGAGCTGCACTCG 1117
Qy 21 -----LeuGlyGlyProGluAlaProCysThrAspProPheGluAlaAlaSer 37
Db 1118 GACTTGGACCTGAGCCGCGCTCCCTGACAGCTCTCG-----CCCTGGAGCCTGCCAC 1171
Qy 38 GlyAlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeuLeu 57
Db 1172 CCCCTGGTCTTGGTGGCAACCTGCTCAGTGGGAAGCTGTCAAGGGGGCTAGCAAGCTG 1231
Qy 58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCys 77
Db -----

```

```

Db 1232 CTCACAGCCCGCAC-----GGCATCGAGTCCCGCGCCAGTTCCTCGCGGAAGAAG 1282
Qy 78 AlaGluValArgGlyPheSerAspLysLysLysLysLysLysLysLysLysLysLysLys 97
Db 1283 GTGCTTCTCCGGGGCTTCTCCCGCAGCACCTCGCGCCTATGCCAGAGGATGTTCCCC 1342
Qy 98 AspGluArgAlaGluArgAlaTyArgPheVallysGluAsnGluThrLeuPheAla 117
Db 1343 GAGCGGGCCCTCGAGGACCGCTCTGTAGCCAGCTGGAGGCCAACCCCACTCTGCGAGC 1462
Qy 118 LeuCysPheValProPheValCysTrpIleValCysThrVal-----LeuArgGln 134
Db 1403 CTGTGCTCTGGCGCCCTCTTCTGTGATCATCTCTCGGCTCTCCAGCACTTCGCTGCT 1462
Qy 135 GlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrSerValTyLeu 154
Db 1463 GCCTTTGAAGGCTCACCACAGCTCCCGACTGCGACGATGACCTCGACAGATGCTCTCTC 1522
Qy 155 LeuPheIleThrSerValLeuSer----- 162
Db 1523 CTG---GTCACTGAGGTCCA-TCTGAACAGGATCGAGCCAGCAGCTGGTGACGCGAA 1578
Qy 163 ---SerAlaProValAlaAspGlyProArgLeuGlnGly-AspLeuArgAsnLeuCysAr 181
Db 1579 CACAGCAGCCAGTGGAGACCTCCACGCGCGCGGACACTCTGTGCTCGCTGGGCGA 1638
Qy 181 gLeuAlaArgGluGlyValLeuGlyArgAlaGlnPheAlaGluLysGluLeuGluG1 201
Db 1639 GGTGGCCACCAGGGGATGGAGAGAGCCTTTGTCTTACCAGGAGGAGGTGCGAGGC 1698
Qy 201 nLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysLysLysLysLys 220
Db 1699 CTCGGGCTCGCAGGAGAGACATGCAGTGGCTCTCTG-----CGGGCTTTCGCGGA 1752
Qy 221 ---GlyValLeuGluThrGluValThrTyGlnPheIleAspGlnSerPheGlnGluPh 239
Db 1753 GCTGGGCCCCGGGGTGACAGCAGTCTATGAGTTTTCACCTCACCTCCAGGCGCTT 1812
Qy 239 eLeuAlaAlaLeuSerTyLeuLeuGluAsp----- 249
Db 1813 CTTTACAGCCTTCTCTCTGCTGGAGCAGAGGTGGGCACTCAGGAGCTGCTCAGGTT 1872
Qy 250 -----GlyGlyValProArgThrAlaAla----- 257
Db 1873 CTTCCAGGAGTGGATGCCCTCGGGGCGAGCAGCAGCTCTCTATCTCTCCCTTCT 1932
Qy 258 -----GlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHis 272
Db 1933 CCCGTTCCAGTCCCTGCGAGGCGAGTGTCCGGCGGGGAAGACCTCTTCAAGAACAGGA 1992
Qy 272 rHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAs 292
Db 1993 TCACCTTCCAGTCAACCAACCTCTCTCTGTCGGGCTGTGTGCCAAAGCCAAACAGAAA-- 2050
Qy 292 pIleGluArgHisPheGlyCysMetValSerGluArgVallysGlnGluAlaLeuArgTr 312
Db 2051 CTCTCGCGCATCTGTGTCCCGGCGGCGCTGAGGAGAAAGCGCAAGGCCCTG---TG 2106
Qy 312 p-----ValGlnGlyGlnGlyCysProGlyValAlaProG1 326
Db 2107 GGCACACCTGTTTCCAGCCTCGGGGCTACTCTGAAGAGCCTGCCCGCTTCAGTCTGA 2166
Qy 326 uValThrGluGlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGluG1 346
Db 2167 AAGCTTCAACAGGTGCGAGGCCATG----- 2191
Qy 346 yGluGluProAsnTyProLeuGluLeuTyCysLeuTyCysLeuTyGluThrGlnGluAsp-- 365
Db 2192 -----CCACAGTTC---ATCTGATGCTCGCTGCTATCTACGAGACACAGAGCCAGAA 2242
Qy 366 -----AlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuArgVa 383
Db 2242 GGTGGGGGAGCTGGGGCGGCGCATCTGCGCC---AATACCTCAAGCTGACCTACTG 2298

```



```
QY 393 lArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCysProAlaG1 403
DB 2299 CAACGCCCTGCTCGCGCGCTCAGCGCGCTCTCCCTGCTGATCACTTCCGCCAAGCG 2358
QY 403 yGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluLysLysLysSe 423
DB 2359 GCTGCGCCCTAGACCTAGACAAACAAATCTC----- 2389
QY 423 tLeuGlyLysArgLeuGlnAlaSerLeuGlyGlySerSerGlnGlyThrThrLysG1 443
DB 2390 -----AAGCACTACCGCGCGGAG-- 2410
QY 443 nLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLe 463
DB 2411 -----CTGACGCCCTGCTTACGCCCTCACT----- 2437
QY 463 uSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerG1 493
DB 2438 ----GTTCTCAGACTCAGCGTAAACAGATCACTACGGTGGGTAAAGTGCTAAGCGA 2493
QY 493 uAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerG1 503
DB 2494 AGAGCTGACCAAAATACAAAATTTGACCTATCTGGGTTTATACAAACAAACAGATCACC 2533
QY 503 uAlaGlyLeuArgMetLeuSerGluGlyLeuAlaIrpProGlnCysArgValGlnThrVa 523
DB 2534 TGTCGAGCCAGGTACGTACCAAAATCTCTG-----GATGAATGCAAGGCTCAGC-- 2635
QY 523 lArgValGlnLeuProAsp-----ProGlnArgGlyLeuGlnTyrLeu----- 537
DB 2606 -CATCTTAAGTGGGNAAAACAAATAACAAGTGAAGGAGGAAGTATCTGCCCTGCG 2664
QY 538 -----ValGlyMet----- 540
DB 2665 TGTGAAGAACAGCAAAATCAATCTCTGAGGTGGGATGTGGGCAATCAAGTTGGGATGA 2724
QY 541 -----LeuArgGlnSerProAlaLeuThrThrLeuAspLe 552
DB 2725 AGGACAAAAGCCTTCGACAGAGGCTCTCGGGAACCAACCCAGCTTGACACCCCTGAGTCT 2764
QY 552 uSerGlyCysGlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuGlnHisG1 572
DB 2785 TCGGTCACAGCATCTCCACAGAGGAGGAGAGGCTTCGAGGGCCCTGCAG--CA 2841
QY 572 nGlyCysGlyLeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuG1 592
DB 2842 GAACACGTCCTAGAAATACTGTGGCTGACCCCAAAATGAATCAACGATSAAGTGGCAGA 2901
QY 592 nGluLeuGlnAlaValLysArgAlaLysProAspLeu-----ValIleThrHisProAl 610
DB 2902 GAGTTGGCAGAAATGTT--GAAAGTCAACACAGACGTTAAGCATTTATGGCTTATCCAGA 2960
QY 610 aLeuAsp 612
DB 2961 ATCAGAT 2967
```

RESULT 12

```
US-09-245-281-7
; Sequence 7, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
; FILE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 37334/118001
; CURRENT APPLICATION NUMBER: US/09/245,281
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: US 09/207,359
; EARLIER FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: US 09/099,041
; EARLIER FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: US 09/019,942
```

; EARLIER FILING DATE: 1998-02-06

```
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3382
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-281-7
```

Alignment Scores: 8.8e-20 Length: 3382
Pred. No.: 172 Matches: 172
Score: 289.00 Conservative: 94
Percent Similarity: 37.84%
Best Local Similarity: 24.47% Mismatches: 281
Query Match: 8.98% Indels: 158
DB: 27 Gaps: 27

US-10-029-347-2 (1-625) x US-09-245-281-7 (1-3382)

```
QY 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
DB 1058 CTGCTGCGCTTCCGCCACGTGGGCCCTTTCACCTTCGATGGCTGGACGAGCTGCACCTG 1117
QY 21 -----LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSer 37
DB 1118 GACTTGGACCTGAGCGCGCTGCTGACAGCTCTGCG-----CCCTGGAGCCTGCCAC 1171
QY 38 GlyAlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeu 57
DB 1172 CCCTGCTCTTGTGGCCAACTGCTCAGTGGGAAGCTGTCAAGGGGGCTAGCAAGCTG 1231
QY 58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCys 77
DB 1232 CTCACAGCCGCGACA-----GGCATCGAGTCCCGCGCCAGTTCCTGCGGAAGAAG 1282
QY 76 AlaGluValArgGlyPheSerAspLysAspLysLysLysTyrPheTyrLysPhePheArg 97
DB 1283 GTGCTTCTCGGGGCTTCTCCGCCACCCACTCGCGGCCCTATGCCAGGAGGATGTCCCC 1342
QY 98 AspGluArgArgAlaGluArgAlaTyrArgPheValLysGluAsnGluThrLeuPheAla 117
DB 1343 GAGCGGCGCTGCGAGGACCCCTGCTGAGCAGCTGGAGGCCAACCCCAACCTCTGCGAGC 1402
QY 118 LeuCysPheValProPheValCysTrpIleValCysThrVal-----LeuArgGln 134
DB 1403 CTGTGCTCTGCGCCCTTCTCTGCTGATCATCTTCGGTGTCTCCAGCACTTCCGTGCT 1462
QY 135 GlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValTyrLeu 154
DB 1463 GCCTTTGAAGGCTCACACAGCTGCCGAGCTGCACGATGACCTGCACAGATGTCTTCTC 1522
QY 155 LeuPheIleThrSerValLeuSer----- 162
DB 1523 CTG---GTCACTGAGGTCCA-TCTGAACAGGATGCAGCCCGAGCGCTGTGCGAGCGAA 1578
QY 163 ---SerAlaProValAlaAspGlyProArgLeuGlnGly-AspLeuArgAsnLeuCysAr 181
DB 1579 CACAGCAGCCAGTGGAGACCTCCACGCGCGCGGACACTCTGTGCTCGCTGGCGGCA 1638
QY 181 gLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGluG1 201
DB 1639 GGTGGCCACCGGGGCGATGAGAAGAGCCCTTTGTCTTCCACCAGGAGGAGGTGCAGGC 1698
QY 201 nLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro-- 220
DB 1699 CTCCGGCTGCAGGAGACACATGCACTGGGCTTCTCTG-----CGGCTTTGCCGGA 1752
QY 221 ---GlyValLeuGluThrGluValThrTyrGlnPheIleAspGlnSerPheGlnGluPh 239
DB 1753 GCTGGGCGCGGGGTGACCAGCAGTCTCTATGAGTTTTTTCACCTCCTCACCGCTT 1812
QY 239 eLeuAlaLeuSerTyrLeuLeuGluAsp----- 249
```

```
Db 1813 CTTTACAGCCTTCTTCTGCTGTGTSAGACAGAGGGTGGGCACTCAGAGCTGCTCAGGTT 1872
Qy 250 -----G-yGlyValProArgThrAlaala 257
Db 1873 CTTCCAGAGTGGATGCCCTCGGGGGGAGCACCACGCTCTGCTATCTCCCTTCCT 1932
Qy 258 -----G-yGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHis 272
Db 1933 CCGCTTCAGTGCCTCGAGGAGCAGTGTCTCCGGCGCGGAGAGACCTCTTCAAGAACACAGGA 1992
Qy 272 rHisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAs 232
Db 1993 TCACCTCCAGTTCACCAACCTCTTCTGCTGGGCTGTGTCCAAAGCCAAACAGAAA-- 2050
Qy 292 pIleGluArgHisPheGlyCysMetValSerGluArgValGlyGlnGluAlaLeuArgTr 312
Db 2051 -CTCCTGGCGCATCTGTGTGTCGCGGAGCCCTCAGGAGAAAGCGCAAGGCCCTG--TG 2106
Qy 312 p-----ValGlnGlyGlnGlyGlnGlyCysProGlyValAlaProG 326
Db 2107 GGCACACCTGTTTCCAGCCTGCGGGGCTACCTGAAGAGCCTGCCCGCTTCAGGTGGA 2166
Qy 326 uValThrGluGlyAlaGlyGlyLeuGluAspThrGluGluProGluGluGluGluG 346
Db 2167 AAGCTTCAACAGGTGCAGGCCATG----- 2191
Qy 346 yGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyrGluThrGlnGluAsp-- 365
Db 2192 -----CCCAGGTTTC--ATCTGATGCTCGGTCATCTACGAGACACAGAGCCAGAA 2241
Qy 366 -----AlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgVa 383
Db 2242 GGTGGGCGAGTGGGGCGAGGCGATCTGGCC--AACTACCTCAAGCTGACCTACTG 2298
Qy 383 lArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCysProAlaGi 403
Db 2299 CAACCCCTGCTCGGGCGAGCTGCGGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2359
Qy 403 yGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluGlyGlyGlySse 423
Db 2359 GCTGCGCTTAGACCTAGACCTAGACCAACATCTC----- 2389
Qy 423 tLeuGlyLysArgLeuG:nA--aserLeuGlyGlyGlySerGlnGlyThrThrLysGl 443
Db 2390 -----AAGCACTACGGCGCGGAG-- 2410
Qy 443 nLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLe 463
Db 2411 -----C-GCAGCCTGCTTCAGCGGCTCAGCT----- 2437
Qy 463 uSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerGl 483
Db 2438 ---GTTCTCAGACTCAGCGTAACACAGATCAGTACGCGTGGGTAAAGGTGTAAACCGA 2493
Qy 483 uAlaLeuArgAlaAlaProAlaLeuThrGluLeuLeuHisAsnArgLeuSerGl 503
Db 2494 AGAGCTGACCAATAACAAATTTGTGACCTATTGTGGGTTTATACAAACACAGATCACCGA 2553
Qy 503 :AlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrVa 523
Db 2554 TGTGGAGCCAGGTAGTCACCAAAATCCTG-----CATGAATGCAAAGGCGCTCAGC-- 2605
Qy 523 lArgValGlnLeuProAsp-----ProGlnArgGlyLeuGlnTyrLeu----- 537
Db 2606 -CATCTTAACTGGGAAATAACAAATAACAAAGTAAGGAGGAAGTATCTCGCCCTGGC 2664
Qy 538 -----ValGlyMet----- 540
Db 2665 TGTGAAGAACAGCAAAATCAATCTCTGAGGTGGGATGTGGGGCAATCAAGTTGGGGATGA 2724
Qy 541 -----LeuArgGlnSerProAlaLeuThrThrThrThrThrThrThrThrThrThr 552
Db 2725 AGGACAAAGCCTTCGGAGAGGCTCTGGGNAACCCAGCCAGCTTGACCCCTTGAGTCT 2784
```

```
Qy 552 uSerGlyCysGlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuGlnHisGl 572
Db 2785 TGGCTCCACAGGCACTCCACAGAGGAGGAGGCTTGGGAGGCCCTGCAG--CA 2841
Qy 572 nGlyCysGlyLeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuG 592
Db 2842 GAACACGCTCTAGAAATACTGTGGCTGACCCAAATGAATCAACGATGAAGTGGCAGA 2901
Qy 592 nGluLeuGlnAlaValLysArgAlaLysProAspLeu-----ValIleThrHisProAl 610
Db 2902 GAGTTTGGCAGAAATGT-T-GAAAGTCAACAGAGGTTAAAGCATTTATGCTTATCCAGA 2960
Qy 610 aLeuAsp 612
Db 2961 ATCAGAT 2967

RESULT 13
US-09-207-359B-7
; Sequence 7, Application US/09207359B
; Patent No. 6469140
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-112001
; CURRENT APPLICATION NUMBER: US/09/207,359B
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (245)...(3103)
US-C9-207-359B-7

Alignment Scores:
Pred. No.: 8,8e-20 Length: 3382
Score: 289.00 Matches: 172
Percent Similarity: 37.84% Conservative: 94
Best Local Similarity: 24.47% Mismatches: 281
Query Match: 8.98% Indels: 158
DB: 4 Gaps: 27

US-10-029-347-2 (1-625) x US-09-207-359B-7 (1-3382)
Qy 1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
Db 1058 CTGCTGGCTTCCGCCAGTCGGGCTCTTACCTTCGATGGCTGGAGAGCTGCATCTG 1117
Qy 21 -----LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSer 37
Db 1118 GACTTGGACCTGAGCGCGTGCCTGACAGCTCTCTG-----CCCTGGAGCCTGCCAC 1171
Qy 38 GlyAlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeu 57
Db 1172 CCCTGCTGCTTGTGGCAACCTGCTCAGTGGGAAGCTGCTCAAGGGGCTAGCAAGCTG 1231
Qy 58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlyCys 77
Db 1232 CTCACAGCCCGCACA-----GGCATCGAGTCCCGCCAGTCTCTCGCGGAAGAG 1282
Qy 78 AlaGluValArgGlyPheSerAspLysLysLysLysLysLysLysLysLysLysLys 97
Db 1283 GTGCTTCTCGGGGCTTCTCCCGCAGCACCTGCGCGCTATGCCAGGAGGATGTTCCCC 1342
```


; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (245)...(3104)
 US-09-340-620A-7

Alignment Scores:

Pred. No.: 8,8e-20 Length: 3382
 Score: 289.00 Matches: 172
 Percent Similarity: 37.84% Conservative: 94
 Best Local Similarity: 24.47% Mismatches: 258
 Query Match: 8.98% Indels: 151
 DB: 4 Gaps: 27

US-10-029-347-2 (1-625) x US-09-340-620A-7 (1-3382)

Qy 1 MetLeuAlaGlnProGlnArgLeuPheLeuLeuAspGlyAlaAspGluLeuProAla 20
 Db 1058 CTGCTGGCTTCCCGCCAGCTGGCTTTCACCTTCGATGGCTGGAGCTGCACCTCG 1117
 Qy 21 -----LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSer 37
 Db 1118 GACTTGGACCTGAGCGCGGT3CCTGACAGCTCCCTGC-----CCCTGGAGCCTGCCAC 1171
 Qy 38 GlyAlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeu 57
 Db 1172 CCCCTGGTCTTGCTGGCCAACTCTCAGTGGGAAGCTGCTCAGAGGGGCTAGCAAGCTG 1231
 Qy 58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuCysSerProGlnCys 77
 Db 1232 CTCACAGCCCGCCACA-----GGCATCAGAGTCCCGGCCAGTTCCTCGCGAAGAG 1282
 Qy 78 AlaGluValArgGlyPheSerAspLysAspLysLysLysLysLysLysPheTyrLysPheArg 97
 Db 1283 GTGCTTCTCCGGGGTCTTCCCGCAGCCACTGCGCCCTATGCCAGAGGATGTTCCCC 1342
 Qy 98 AspGluArgAlaGalaGalaAlaTyrArgPheValLysGluAsnGluThrLeuPheAla 117
 Db 1343 GAGCGGGCCCTGCAGGACCGCTCTGAGCCAGCTGGAGGCCAACCCCAACCTCTGCAGC 1402
 Qy 118 LeuCysPheValProPheValCysTrpLeuValCysThrVal-----LeuArgGln 134
 Db 1403 CTGCGCTCTGCGCCCTCTTCTGTGGATCATCTCCGGTGCCTCCAGACATCCCGTCT 1462
 Qy 135 GlnLeuGluLeuGlyArgAspLeuSerArgThrSerLysThrThrThrSerValThrLeu 154
 Db 1463 GCCTTGAAGGCTCACCACAGCTGCCGAGCTGCACGATGACCTGTGACAGATGTTCTCTC 1522
 Qy 155 LeuPheLeuThrSerValLeuSer-----LeuSer----- 162
 Db 1523 CTG---GTCACTGAGGTCCA-TCTGAACAGGATGCAGCCAGCCCTGTGCGAGCGAA 1578
 Qy 163 ---SerAlaProValAlaAspGlyProArgLeuGlnGly-AspLeuArgAsnLeuCysAr 181
 Db 1579 CACAGGACCCCTGAGTGAGACCTCCAGCGCGCGGGAGACACTCTGCTCGCTGGGSCA 1639
 Qy 181 gLeuAlaArgGluGlyValLeuGlyArgArgAlaGlnPheAlaGluLysGluLeuGlu 201
 Db 1639 GGTGCCCCACCGGGCATGGAGAAGAGCTCTTGTCTTCCACCCAGGAGGAGGTGCAGGC 1698
 Qy 201 nLeuGluLeuArgGlySerLysValGlnThrLeuPheLeuSerLysLysGluLeuPro 220
 Db 1699 CTCGGGCTGAGGAGAGACATGCAGCTGGCTTCTCTG-----CGGGCTTTCGCGGA 1752
 Qy 221 ---GlyValLeuGluThrGluValThrThrGlnPheLeuAspGlnSerPheGlnGluPh 239
 Db 1753 GCTGGCCCCGGGGTGCAGCAGCAGCTCTATGAGTGTTCACCTCACCCTCCAGGCTT 1812
 Qy 239 eLeuAlaAlaLeuSerTyrLeuLeuLeuAsp----- 249
 Db 1813 CTTACAGCCTTCTTCTCGTGTGGACGACAGGGTGGGCACTCAGGAGCTGCTCAGGTT 1872

Qy 250 -----GlyGlyValProArgThrAlaAla----- 257
 Db 1873 CTTCCAGGAGTGGATGCCCCCTGGGGGCGAGCACCACGTCCTGCTATCCTCCCTTCTCT 1932
 Qy 258 -----GlyGlyValGlyThrLeuLeuArgGlyAspAlaGlnProHis 272
 Db 1933 CCCGTTCAGTCCCTGCGAGGCGAGTGGTCCGGCGCGGAAGACCTCTTCAAGAACAAGGA 1992
 Qy 272 HisLeuValLeuThrThrArgPheLeuPheGlyLeuLeuSerAlaGluArgMetArgAs 292
 Db 1993 TCACCTCCAGTTCCACCAACCTCTTCTCTGCGGGCTGTGTGCTCAAAGCCCAACAGAAA-- 2050
 Qy 292 PileGluArgHisPheGlyCysMetValSerGluArgValLysGlnGluAlaLeuArgTr 312
 Db 2051 -CTCTCGGCGATCTGTGTCGGCGGCGAGCCCTGAGGAGAAGCGCAAGGCCCTG---TG 2106
 Qy 312 p-----ValGlnGlyGlnGlyGlnGlyCysProGlyValAlaProGln 326
 Db 2107 GGCACACCTGTTTCCAGCCTGCGGGCTACCTGAAGAGCCTGCCCGCGTTCAGGTCTGA 2166
 Qy 326 uValThrGluGlyAlaLysGlyLeuGluAspThrGluGluProGluGluGluGlu 346
 Db 2167 AGCTTCAACAGGTGCGAGCCATG----- 2191
 Qy 346 yGluGluProAsnTyrProLeuGluLeuLeuTyrCysLeuTyrGluThrGlnGluAsp 365
 Db 2192 -----CCACAGTTC---ATCTGATGCTGCGCTGCTATCTACGACACAGAGCCAGAA 2241
 Qy 366 -----AlaPheValArgGlnAlaLeuCysArgPheProGluLeuAlaLeuGlnArgVa 383
 Db 2242 GGTGGGCGAGCTGGCGGCGAGGCATCTGCGCC---AACTACCTCAAGCTGACCTACTG 2298
 Qy 383 ArgPheCysArgMetAspValAlaValLeuSerTyrCysValArgCysCysProAlaGln 403
 Db 2299 CAACGCTGCTCGCGCCGACTGCGAGGCCCTCTCTCTGCTGCTGCTCCTCTCCCAAGCG 2358
 Qy 403 yGlnAlaLeuArgLeuLeuSerCysArgLeuValAlaAlaGlnGluLysLysLysSe 423
 Db 2359 GCTGGCCCTAGACCTAGACCAACAATCTC----- 2389
 Qy 423 rLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGlyThrThrLysGln 443
 Db 2390 -----AACGACTACGGCGTGGCGAG-- 2410
 Qy 443 nLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLe 463
 Db 2411 -----CTGCAGCCCTGCTTCCAGCCGCTCACT----- 2437
 Qy 463 uSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerGln 483
 Db 2438 ----GTTCTCAGACTCAGCGTAAACACAGATCAGCGGTGGGTAAAGGTGCTTAAGCGA 2493
 Qy 483 uAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerGln 503
 Db 2494 AGAGCTGACCAATACAAAATTGAGCTATTGGGTATTATACAAACACAGATCACCAGA 2553
 Qy 503 uAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrVa 523
 Db 2554 TGTGCGAGCCAGGTACGTACCAAAAACCTG-----GATGAATGCAAAAGCCCTCAGC-- 2605
 Qy 523 lArgValGlnLeuProAsp-----ProGlnArgGlyLeuGlnThrLeu 537
 Db 2606 -CATCTTAACTGGGAAAAAACAATAACAAAGTAAGAGGAGGAATATCTCGCCCTGGC 2664
 Qy 538 -----ValGlyMet----- 540
 Db 2665 TGTGNAGAACAGCAATCAATCTCTGAGGTGGATGTGGGGCAATCAAGTTGGGGATGA 2724
 Qy 541 -----LeuArgGlnSerProAlaLeuThrThrLeuAspLe 552
 Db 2725 AGGACAAAAGCCTTGGCAGAGGCTCTGCGGAACCAACCCAGCTGTGACCACTGAGTCT 2784
 Qy 552 userGlyCysGlnLeuProAlaProMetValThrThrLeuCysAlaValLeuGlnHisGln 572

```

Db      2785  TGGGTCCAAGCGCATCTCCACAGAAAGAGAGCGCTTGGCAGGGCCCTGCAG---CA 2841
Qy      572  nGlyCysGlyLeuGlnThrIleuSerValGluLeuSerGluGlnSerLeuG1 592
Db      2842  GACACAGTCTCTAGAAATATGTGGCTGAGCCCAAAATGAACTCAACGATGAAGTGGCAGA 2901
Qy      592  nGluLeuGlnAlaValLysArgAlaLysProAspLeu-----ValIleThrHisProAl 610
Db      2902  GAGTTTGGCAGAAATGTT-GAAAGTCAACACAGAGTTTAAAGCATTTATGCTTATCCAGA 2960
Qy      610  aLeuAsp 612
Db      2961  ATCAGAT 2967

RESULT 15
US-09-245-281-38
; Sequence 38, Application: US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 07334/118001
; CURRENT APPLICATION NUMBER: US/09/245,281
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: US 09/207,359
; EARLIER FILING DATE: 1998-12-06
; EARLIER APPLICATION NUMBER: JS 09/099,041
; EARLIER FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: US 09/019,942
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 4302
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (438)....(1284)
; NAME/KEY: misc_feature
; LOCATION: (1)....(4302)
; OTHER INFORMATION: n = A,T,C or G
US-09-245-281-38

Alignment Scores:
Pred. No.: 3,46e-19 Length: 4302
Score: 285.00 Matches: 169
Percent Similarity: 37.66% Conservative: 92
Best Local Similarity: 24.39% Mismatches: 278
Query Match: 8.86% Indels: 156
DB: 4 Gaps: 26

US-10-029-347-2 (1-625) x US-09-245-281-38 (1-4302)
Qy      1 MetLeuAlaGlnProGlnArgLeuLeuPheIleLeuAspGlyAlaAspGluLeuProAla 20
Db      1115  CTGCTGGCTTCCCCACGTTGGCCCTCTTACCCTTCATGGCTGGACGAGCTGCACTCG 1174
Qy      21 -----LeuGlyGlyProGluAlaAlaProCysThrAspProPheGluAlaAlaSer 37
Db      1175  GACTTGACCTGAGCCGGTGCCTGCACAGCTCCTGC-----CCCTGGGAGCCTGCCAC 1228
Qy      38 GAlaArgValLeuGlyGlyLeuLeuSerLysAlaLeuLeuProThrAlaLeuLeu 57
Db      1229  CCCTTGCTCTGTGTGGCAACCTGCTCAGTGGGAAGCTGCTCAAGGGGGGTAGCAAGCTG 1288
Qy      58 ValThrThrArgAlaAlaAlaProGlyArgLeuGlnGlyArgLeuLeuCysSerProGlnCys 77
Db      1289  CTACAGCCCGCACA-----GGCATCGAGGTCCCGCCGACAGTTCCTTCGCGAAGAG 1339
Qy      78 AlGluValArgGlyPheSerAspLysAspLysLysLysThrPheTyIlyLysPhePheArg 97

```

```
Db 2356 CAAGCGCTGCTCGGCGACTGACGGCGCTCTCTTCGTCCTGCATCACTTCCCCAAAGCG 2415
Qy 403 yGlnAlaLeuArgLeuIleSerCysArgLeuValAlaGlnGlnLysLysLysSe 423
Db 2416 GCTGGCCCTAGACCTAGACCAACAATCTC----- 2445
Qy 423 rLeuGlyLysArgLeuGlnAlaSerLeuGlyGlyGlySerSerGlnGlyThrThrLysG1 443
Db 2447 -----AAGCACTACGGCGTGGGAG-- 2467
Qy 443 nLeuProAlaSerLeuLeuHisProLeuPheGlnAlaMetThrAspProLeuCysHisLe 463
Db 2468 -----CTGCAGCCCTGCTTCAGCGCCTCACT----- 2494
Qy 463 uSerSerLeuThrLeuSerHisCysLysLeuProAspAlaValCysArgAspLeuSerG1 483
Db 2495 ----GTTCTCAGACTCAGCGTAAACCAGATCACTGACGGTGGGGTAAAGGTGCTAAGCGA 2550
Qy 483 uAlaLeuArgAlaAlaProAlaLeuThrGluLeuGlyLeuLeuHisAsnArgLeuSerG1 503
Db 2551 AGACCTGACCAAAATACAAAATGTGACCTATTGGGTTTATACAAACAGATCACCSA 2610
Qy 503 uAlaGlyLeuArgMetLeuSerGluGlyLeuAlaTrpProGlnCysArgValGlnThrVa 523
Db 2611 TGTGGAGCCAGGTACGTCAACAAATCTCTG-----GATGAATGCAAGGCCCTCACG-- 2662
Qy 523 lArgValGlnLeuProAsp-----ProGlnArgGlyLeuGlnTyrLeu----- 537
Db 2663 -CATCTTAAACTGGGAAAAAACAATAACAAGTGAAGGAGGAGGAGTATCTCGCCCTGGC 2721
Qy 538 -----ValGlyMet----- 540
Db 2722 TGTGAACAACAGCAATCAATCTCTGAGGTTGGGATGTGGGGCAATCAAGTTGGGATGA 2781
Qy 541 -----LeuArgGlnSerProAlaLeuThrThrLeuAspLe 552
Db 2782 AGGAGCAAAAGCCCTTCGAGAGGCTCTGGCGAACCACCCAGCTTGACCACCTGAGTCT 2841
Qy 552 uSerGlyCysGlnLeuProAlaProMetValThrTyrLeuCysAlaValLeuGlnHisG1 572
Db 2842 TGCCTCAACGGCATCTCCACAGAGGAGGAGAGAGCCTTGGCGAGGCCCTGCAG---CA 2898
Qy 572 nGlyCysGlyLeuGlnThrLeuSerLeuAlaSerValGluLeuSerGluGlnSerLeuG1 592
Db 2899 GAACACGCTCTAGAAAATCTGTGGCTGACCCAAAATGAATCAACGATGAAGTGGCAGA 2958
Qy 592 rGluLeuGlnAlaValLysArgAlaLysProAspLeu 604
Db 2959 GAGTGTGGCAGAAATGTT-GAAAGTCAACCAGACGTT 2994
```

Search completed: October 3, 2003, 12:09:07
Job time : 142 secs